

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2003, 13:07:05 ; Search time 20 Seconds
(without alignments)
1840.522 Million cell updates/sec

Title: US-09-970-287-2
Perfect score: 4751
Sequence: 1 MGPSSIVLCLLSATVFSLLG.....PEMKRPSKSLGQLWEGWEG 870

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep: *
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep: *
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep: *
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep: *
5: /cgn2_6/prodata/1/iaa/PCUS_COMB.pep: *
6: /cgn2_6/prodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4719.5	99.3	867	4	US-09-996-243-84
2	3776	79.5	690	4	US-09-886-319A-36
3	3175.5	66.8	871	4	US-09-773-426A-3
4	3117.5	65.6	867	4	US-09-668-673B-2
5	3059	64.4	1611	4	US-09-668-673B-16
6	2093	44.1	455	4	US-09-668-673B-14
7	1702	35.8	470	4	US-09-886-319A-35
8	1606.5	33.8	360	4	US-09-810-347-4
9	1590.5	33.5	372	4	US-09-810-347-2
10	1589	33.4	709	4	US-09-668-673B-3
11	1490	31.4	309	4	US-09-810-347-5
12	1183	24.9	307	4	US-09-810-347-6
13	827.5	17.4	510	1	US-08-484-493-11
14	827.5	17.4	510	1	US-08-484-494-11
15	827.5	17.4	510	2	US-08-345-212-11
16	827.5	17.4	510	3	US-09-249-003-11
17	827.5	17.4	510	4	US-09-685-844-11
18	565	11.9	160	4	US-09-668-673B-7
19	507	10.7	98	4	US-09-668-673B-22
20	405	8.5	96	4	US-09-668-673B-20
21	399.5	8.4	196	4	US-09-668-673B-18
22	370.5	7.8	520	4	US-09-773-426A-10
23	325	6.8	552	4	US-09-773-426A-9
24	286.5	6.0	406	4	US-09-227-357-187
25	277.5	5.8	556	1	US-08-445-586-2
26	276.5	5.8	554	1	US-08-445-586-7
27	238.5	5.0	583	1	US-08-445-586-9

28	238.5	5.0	583	4	US-09-646-683-1
29	234.5	4.9	550	1	US-08-484-493-2
30	234.5	4.9	550	1	US-08-484-494-2
31	234.5	4.9	550	2	US-08-345-212-2
32	234.5	4.9	550	3	US-09-249-003-2
33	234.5	4.9	550	4	US-09-685-844-2
34	232.5	4.9	583	1	US-08-484-493-14
35	232.5	4.9	583	1	US-08-484-494-14
36	232.5	4.9	583	2	US-08-345-212-14
37	232.5	4.9	583	3	US-09-249-003-14
38	232.5	4.9	583	4	US-09-685-844-14
39	230	4.8	547	4	US-09-252-991A-25705
40	215	4.5	551	1	US-08-484-493-15
41	215	4.5	551	1	US-08-484-494-15
42	215	4.5	551	2	US-08-345-212-15
43	215	4.5	551	3	US-09-249-003-15
44	215	4.5	551	4	US-09-685-844-15
45	208.5	4.4	533	1	US-08-445-586-10

ALIGNMENTS

RESULT 1

US-09-996-243-84
; Sequence 84, Application US/09996243
; Patent No. 6478825
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C13
; CURRENT APPLICATION NUMBER: US/09/996,243
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28

1 ; PRIOR APPLICATION NUMBER: 60/084600
2 ; PRIOR FILING DATE: 1998-05-07
3 ; PRIOR APPLICATION NUMBER: 60/087106
4 ; PRIOR FILING DATE: 1998-05-28
5 ; PRIOR APPLICATION NUMBER: 60/087607
6 ; PRIOR FILING DATE: 1998-06-02
7 ; PRIOR APPLICATION NUMBER: 60/087609
8 ; PRIOR FILING DATE: 1998-06-02
9 ; PRIOR APPLICATION NUMBER: 60/087759
10 ; PRIOR FILING DATE: 1998-06-02
11 ; PRIOR APPLICATION NUMBER: 60/087827
12 ; PRIOR FILING DATE: 1998-06-03
13 ; PRIOR APPLICATION NUMBER: 60/088021
14 ; PRIOR FILING DATE: 1998-06-04
15 ; PRIOR APPLICATION NUMBER: 60/088025
16 ; PRIOR FILING DATE: 1998-06-04
17 ; PRIOR APPLICATION NUMBER: 60/088026
18 ; PRIOR FILING DATE: 1998-06-04
19 ; PRIOR APPLICATION NUMBER: 60/088028
20 ; PRIOR FILING DATE: 1998-06-04
21 ; PRIOR APPLICATION NUMBER: 60/088029
22 ; PRIOR FILING DATE: 1998-06-04
23 ; PRIOR APPLICATION NUMBER: 60/088030
24 ; PRIOR FILING DATE: 1998-06-04
25 ; PRIOR APPLICATION NUMBER: 60/088033
26 ; PRIOR FILING DATE: 1998-06-04
27 ; PRIOR APPLICATION NUMBER: 60/088326
28 ; PRIOR FILING DATE: 1998-06-04
29 ; PRIOR APPLICATION NUMBER: 60/088167
30 ; PRIOR FILING DATE: 1998-06-05
31 ; PRIOR APPLICATION NUMBER: 60/088202
32 ; PRIOR FILING DATE: 1998-06-05
33 ; PRIOR APPLICATION NUMBER: 60/088212
34 ; PRIOR FILING DATE: 1998-06-05
35 ; PRIOR APPLICATION NUMBER: 60/088217
36 ; PRIOR FILING DATE: 1998-06-05
37 ; PRIOR APPLICATION NUMBER: 60/088655
38 ; PRIOR FILING DATE: 1998-06-09
39 ; PRIOR APPLICATION NUMBER: 60/088734
40 ; PRIOR FILING DATE: 1998-06-10
41 ; PRIOR APPLICATION NUMBER: 60/088738
42 ; PRIOR FILING DATE: 1998-06-10
43 ; PRIOR APPLICATION NUMBER: 60/088742
44 ; PRIOR FILING DATE: 1998-06-10
45 ; PRIOR APPLICATION NUMBER: 60/088810
46 ; PRIOR FILING DATE: 1998-06-10
47 ; PRIOR APPLICATION NUMBER: 60/088824
48 ; PRIOR FILING DATE: 1998-06-10
49 ; PRIOR APPLICATION NUMBER: 60/088826
50 ; PRIOR FILING DATE: 1998-06-10
51 ; PRIOR APPLICATION NUMBER: 60/088858
52 ; PRIOR FILING DATE: 1998-06-11
53 ; PRIOR APPLICATION NUMBER: 60/088861
54 ; PRIOR FILING DATE: 1998-06-11
55 ; PRIOR APPLICATION NUMBER: 60/088876
56 ; PRIOR FILING DATE: 1998-06-11
57 ; PRIOR APPLICATION NUMBER: 60/089105
58 ; PRIOR FILING DATE: 1998-06-12
59 ; PRIOR APPLICATION NUMBER: 60/089440
60 ; PRIOR FILING DATE: 1998-06-16
61 ; PRIOR APPLICATION NUMBER: 60/089512
62 ; PRIOR FILING DATE: 1998-06-16
63 ; PRIOR APPLICATION NUMBER: 60/089514
64 ; PRIOR FILING DATE: 1998-06-16
65 ; PRIOR APPLICATION NUMBER: 60/089532
66 ; PRIOR FILING DATE: 1998-06-17
67 ; PRIOR APPLICATION NUMBER: 60/089538
68 ; PRIOR FILING DATE: 1998-06-17
69 ; PRIOR APPLICATION NUMBER: 60/089598
70 ; PRIOR FILING DATE: 1998-06-17
71 ; PRIOR APPLICATION NUMBER: 60/089599
72 ; PRIOR FILING DATE: 1998-06-17
73 ; PRIOR APPLICATION NUMBER: 60/089600
74 ; PRIOR FILING DATE: 1998-06-17
75 ; PRIOR APPLICATION NUMBER: 60/089653
76 ; PRIOR FILING DATE: 1998-06-17
77 ; PRIOR APPLICATION NUMBER: 60/089801
78 ; PRIOR FILING DATE: 1998-06-18
79 ; PRIOR APPLICATION NUMBER: 60/089907
80 ; PRIOR FILING DATE: 1998-06-18
81 ; PRIOR APPLICATION NUMBER: 60/089908
82 ; PRIOR FILING DATE: 1998-06-18
83 ; PRIOR APPLICATION NUMBER: 60/089947
84 ; PRIOR FILING DATE: 1998-06-19
85 ; PRIOR APPLICATION NUMBER: 60/089948
86 ; PRIOR FILING DATE: 1998-06-19
87 ; PRIOR APPLICATION NUMBER: 60/089952
88 ; PRIOR FILING DATE: 1998-06-19
89 ; PRIOR APPLICATION NUMBER: 60/090246
90 ; PRIOR FILING DATE: 1998-06-22
91 ; PRIOR APPLICATION NUMBER: 60/090252
92 ; PRIOR FILING DATE: 1998-06-22
93 ; PRIOR APPLICATION NUMBER: 60/090254
94 ; PRIOR FILING DATE: 1998-06-22
95 ; PRIOR APPLICATION NUMBER: 60/090349
96 ; PRIOR FILING DATE: 1998-06-23
97 ; PRIOR APPLICATION NUMBER: 60/090355
98 ; PRIOR FILING DATE: 1998-06-23
99 ; PRIOR APPLICATION NUMBER: 60/090429
100 ; PRIOR FILING DATE: 1998-06-24
101 ; PRIOR APPLICATION NUMBER: 60/090431
102 ; PRIOR FILING DATE: 1998-06-24
103 ; PRIOR APPLICATION NUMBER: 60/090435
104 ; PRIOR FILING DATE: 1998-06-24
105 ; PRIOR APPLICATION NUMBER: 60/090444
106 ; PRIOR FILING DATE: 1998-06-24
107 ; PRIOR APPLICATION NUMBER: 60/090445
108 ; PRIOR FILING DATE: 1998-06-24
109 ; PRIOR APPLICATION NUMBER: 60/090472
110 ; PRIOR FILING DATE: 1998-06-24
111 ; PRIOR APPLICATION NUMBER: 60/090535
112 ; PRIOR FILING DATE: 1998-06-24
113 ; PRIOR APPLICATION NUMBER: 60/090540
114 ; PRIOR FILING DATE: 1998-06-24
115 ; PRIOR APPLICATION NUMBER: 60/090542
116 ; PRIOR FILING DATE: 1998-06-24
117 ; PRIOR APPLICATION NUMBER: 60/090557
118 ; PRIOR FILING DATE: 1998-06-24
119 ; PRIOR APPLICATION NUMBER: 60/090676
120 ; PRIOR FILING DATE: 1998-06-25
121 ; PRIOR APPLICATION NUMBER: 60/090678
122 ; PRIOR FILING DATE: 1998-06-25
123 ; PRIOR APPLICATION NUMBER: 60/090690
124 ; PRIOR FILING DATE: 1998-06-25
125 ; PRIOR APPLICATION NUMBER: 60/090694
126 ; PRIOR FILING DATE: 1998-06-25
127 ; PRIOR APPLICATION NUMBER: 60/090695
128 ; PRIOR FILING DATE: 1998-06-25
129 ; PRIOR APPLICATION NUMBER: 60/090696
130 ; PRIOR FILING DATE: 1998-06-25
131 ; PRIOR APPLICATION NUMBER: 60/090862
132 ; PRIOR FILING DATE: 1998-06-26
133 ; PRIOR APPLICATION NUMBER: 60/090863
134 ; PRIOR FILING DATE: 1998-06-26
135 ; PRIOR APPLICATION NUMBER: 60/091360
136 ; PRIOR FILING DATE: 1998-07-01
137 ; PRIOR APPLICATION NUMBER: 60/091478
138 ; PRIOR FILING DATE: 1998-07-02
139 ; PRIOR APPLICATION NUMBER: 60/091544
140 ; PRIOR FILING DATE: 1998-07-01
141 ; PRIOR APPLICATION NUMBER: 60/091519
142 ; PRIOR FILING DATE: 1998-07-02
143 ; PRIOR APPLICATION NUMBER: 60/091626
144 ; PRIOR FILING DATE: 1998-07-02
145 ; PRIOR APPLICATION NUMBER: 60/091633
146 ; PRIOR FILING DATE: 1998-07-02

; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

```
Query Match      99.3%; Score 4719.5; DB 4; Length 867;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 MGPPSLVLCCLLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIILVLTDDQVDELGS 60
DB 1 MGPPSLVLCCLLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIILVLTDDQVDELGS 60

QY 61 MQVMNKTIRIMEQGTHFINAFVTTPMCCPSRSSILTKGYVHNHTYNNENCSSPSWQA 120
DB 61 MQVMNKTIRIMEQGAHFINAFVTTPMCCPSRSSILTKGYVHNHTYNNENCSSPSWQA 120

QY 121 QHESRTEAVILNSTGYRTAFGKYLYNEVSGSVPPGKMWGLLKNRSFYNYTLCRNGVK 180
DB 121 QHESRTEAVILNSTGYRTAFGKYLYNEVSGSVPPGKMWGLLKNRSFYNYTLCRNGVK 180

QY 181 EXHGSYSDKDYLTDLITNDSVSFFRTSKOMYPHRPVLWVISHAAPHGPDAPQYSRLFP 240
DB 181 EXHGSYSDKDYLTDLITNDSVSFFRTSKOMYPHRPVLWVISHAAPHGPDAPQYSRLFP 240

QY 241 NASQHTPSYVAPNPDKHMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
DB 241 NASQHTPSYVAPNPDKHMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300

QY 301 LVETGELONTYIVYTADHGHIHQFGLVKGSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360
DB 301 LVETGELONTYIVYTADHGHIHQFGLVKGSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360

QY 361 LNIDLAPTILDIAGLDIPADMDGKSLILKLDTERPVNRFHLKKQVWRDSFLVERGKLL 420
DB 361 LNIDLAPTILDIAGLDIPADMDGKSLILKLDTERPVNRFHLKKQVWRDSFLVERGKLL 420

QY 421 HKRDNDKYDAQENFLPKYQRYKDLQCORAEYQTACEQLGQKMQCVEDATGKLLHKCKGP 480
DB 421 HKRDNDKYDAQENFLPKYQRYKDLQCORAEYQTACEQLGQKMQCVEDATGKLLHKCKGP 480

QY 481 MRLGGSRALSNLVPKYQGSSEACTCDSGDYKLSLAGRRKKLKKKCYKASVRSRSIRSV 540
DB 481 MRLGGSRALSNLVPKYQGSSEACTCDSGDYKLSLAGRRKKLKKKCYKASVRSRSIRSV 540

QY 541 ALEVGRVYHVLGLDAAQPRNLTKHWPAGPDQDDKGGDFSGTGGLPDYSAANPIKVT 600
DB 541 ALEVGRVYHVLGLDAAQPRNLTKHWPAGPDQDDKGGDFSGTGGLPDYSAANPIKVT 600

QY 601 HRCYILENDTVQCDLDLYKSLQAWDKHLHIDHEIETLQNKIKNLEVRGHLKKRPEEC 660
DB 601 HRCYILENDTVQCDLDLYKSLQAWDKHLHIDHEIETLQNKIKNLEVRGHLKKRPEEC 660

QY 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQEKDKVWLLRQKRRKKLLKRLQNNNDTC 720
DB 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQEKDKVWLLRQKRRKKLLKRLQNNNDTC 720

QY 721 SMPGLTCTFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLECFEATGFLE 780
DB 721 SMPGLTCTFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLECFEATGFLE 780

QY 781 YFDLNTDPQLMNAVNTLDRVNLQHLVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840
DB 781 YFDLNTDPQLMNAVNTLDRVNLQHLVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840

QY 841 QYRQFORRWKPEWKPSSKSLCOLMEGHEG 870
DB 838 QYRQFORRWKPEWKPSSKSLCOLMEGHEG 867
```

RESULT 2

US-09-886-319A-36
; Sequence 36, Application US/09886319A
; Patent No. 6586185
; GENERAL INFORMATION:
; APPLICANT: Wolf, Eckard
; APPLICANT: Werner, Sabine
; APPLICANT: Halle, Jorn-Peter
; APPLICANT: Regenbogen, Johannes
; APPLICANT: Coppel, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
; TITLE OF INVENTION: Active Substances
; FILE REFERENCE: 50125/014002
; CURRENT APPLICATION NUMBER: US/09/886,319A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/222,081
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: DE 10030149.5
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-319A-36

```
Query Match      79.5%; Score 3776; DB 4; Length 690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 EXHGSYSDKDYLTDLITNDSVSFFRTSKOMYPHRPVLWVISHAAPHGPDAPQYSRLFP 240
DB 1 EXHGSYSDKDYLTDLITNDSVSFFRTSKOMYPHRPVLWVISHAAPHGPDAPQYSRLFP 60

QY 241 NASQHTPSYVAPNPDKHMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
DB 61 NASQHTPSYVAPNPDKHMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 120

QY 301 LVETGELONTYIVYTADHGHIHQFGLVKGSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 360
DB 121 LVETGELONTYIVYTADHGHIHQFGLVKGSMPEYFDIRVPFVVRGPNVEAGCLNPHIV 180

QY 361 LNIDLAPTILDIAGLDIPADMDGKSLILKLDTERPVNRFHLKKQVWRDSFLVERGKLL 420
DB 181 LNIDLAPTILDIAGLDIPADMDGKSLILKLDTERPVNRFHLKKQVWRDSFLVERGKLL 240

QY 421 HKRDNDKYDAQENFLPKYQRYKDLQCORAEYQTACEQLGQKMQCVEDATGKLLHKCKGP 480
DB 241 HKRDNDKYDAQENFLPKYQRYKDLQCORAEYQTACEQLGQKMQCVEDATGKLLHKCKGP 300

QY 481 MRLGGSRALSNLVPKYQGSSEACTCDSGDYKLSLAGRRKKLKKKCYKASVRSRSIRSV 540
DB 301 MRLGGSRALSNLVPKYQGSSEACTCDSGDYKLSLAGRRKKLKKKCYKASVRSRSIRSV 360

QY 541 ALEVGRVYHVLGLDAAQPRNLTKHWPAGPDQDDKGGDFSGTGGLPDYSAANPIKVT 600
DB 361 ALEVGRVYHVLGLDAAQPRNLTKHWPAGPDQDDKGGDFSGTGGLPDYSAANPIKVT 420

QY 601 HRCYILENDTVQCDLDLYKSLQAWDKHLHIDHEIETLQNKIKNLEVRGHLKKRPEEC 660
DB 421 HRCYILENDTVQCDLDLYKSLQAWDKHLHIDHEIETLQNKIKNLEVRGHLKKRPEEC 480

QY 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQEKDKVWLLRQKRRKKLLKRLQNNNDTC 720
DB 481 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQEKDKVWLLRQKRRKKLLKRLQNNNDTC 540

QY 721 SMPGLTCTFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLECFEATGFLE 780
DB 541 SMPGLTCTFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLECFEATGFLE 600

QY 781 YFDLNTDPQLMNAVNTLDRVNLQHLVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840
```

Db 601 YFDLNDTPYOLMNAVNTLDROVLNQLRVOLMELRSCKGYKCNPRTRNMDLGLKGGSYE 660
Qy 841 QYRQORRKPMPKPPSSKSLGQWEGWEG 870
Db 661 QYRQORRKPMPKPPSSKSLGQWEGWEG 690

RESULT 3
US-09-773-426A-3
; Sequence 3, Application US/09773426A
; Patent No. 6534302
; GENERAL INFORMATION:
; APPLICANT: Gluckman, Maria Alexandra
; APPLICANT: Williamson, Mark
; APPLICANT: Teia, Fong-Ying
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: 22438, 23553, 25278, and 26212 No. 6534302el
; FILE OF INVENTION: Human Sulfatases (A CIP Application)
; FILE REFERENCE: 35800/208398(5800-79
; CURRENT APPLICATION NUMBER: US/09/773,426A
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 09/495,823
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 871
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-773-426A-3

Query Match 66.8%; Score 3175.5; DB 4; Length 871;
Best Local Similarity 64.9%; Pred. No. 3.1e-301;
Matches 578; Conservative 120; Mismatches 142; Indels 51; Gaps 12;

Qy 5 SLVLCILSATVFSLLGGSSAFSLSHRLKGRFORDRNRIRPNILVLTDDQDVELGSMQVM 64
Db 7 ALVLAVLGT---ELLSGLSTVRSFRGRIOQRKIRPNILVLTDDQDVELGSLQVM 63
Qy 65 NKTRIMEOGGTHFINAFVTTMCCPSRSSILTKYVHNHNTYTNENCSPPSQAQHES 124
Db 64 NKTRIMEGSGATFINAFVTTMCCPSRSSMLTKYVHNHNTYTNENCSPPSQAQHEP 123
Qy 125 RFAVYLNSTGYRTAFPGKYLYNEYSYVPGKEMVGLLKNRSFYNTLCRNGVKEKHG 184
Db 124 RFAVYLNSTGYRTAFPGKYLYNEYSYVPGKEMVGLLKNRSFYNTLCRNGVKEKHG 183
Qy 185 SDYSKDYLTDLTNDVSVFRTSKMYPHRPVLMVISHAAPHGSDSAPQYSLRFPNASQ 244
Db 184 FQYAKDYFTDLITNESINYPKSKMYPHRPVLMVISHAAPHGSDSAPQYSLRFPNASQ 243
Qy 245 HITPSYNYAPNDKHWIMQYTGPMPIHMEFTNMLORKELOTLMSVDDSMETIYNMLVET 304
Db 244 HITPSYNYAPNDKHWIMQYTGPMPIHMEFTNMLORKELOTLMSVDDSMETIYNMLVET 303
Qy 305 GELDNTYIVYVYADHGYHIGQGLVKGSKMYPBFDIRVPYVPGVNPVAGCLNPHIVLNID 364
Db 304 GELENTYIIVYVYADHGYHIGQGLVKGSKMYPBFDIRVPYVPGVNPVAGCLNPHIVLNID 363
Qy 365 LAPTLIDLAGLDIPADMOKSILKLDTRPVNRPHLKKOMVRWDSFIVERGKLLHGRD 424
Db 364 LAPTLIDLAGLDTPPDVGGKSVLKLDPKPGNFRFTNKAIRWDTFLVERGKFLRKE 423
Qy 425 NDKVDAQEEFLPKYORVNDLCORAEVOTACELGOKMOCVEDATGKLLHKKCPMELG 484
Db 424 ESSKNIOQSNHLPKYERVELCQARYQACQPGQKWCOCIEDTSGKLIHKKAGPSDLL 483
Qy 485 GSR-ALSNLPKYVYGGSEACTCDSDGYKLSLAGRR-KKLFFK-----KYKASYVRSR1 537
Db 484 TVRQSTRNLYARGFHDKDKCSKESGYSRASRSQKSRQFLRNQOTPKYKPRFVHTROT 543
Qy 538 RSVAVIEDGVRYHVGLGD-----AAGPRNLTKRHNEG--APEDQDDKGG-----DFS 583

Db 544 RSLSVFEFEYIDINLEBEEBELQVLQPRNIAKRHDEGHKGRDLOQASSGGNRRGRLADSS 603
Qy 584 GTGGGLPDYSAANPIKVTTHRCYILENDTVQCDLDLYKSLQAKWKHKLHDIETLQNKIK 643
Db 604 NAVGPP-----TTVRVTHKCFILPNDISHCBRELYQSARAWKDKAYIDKEIEALQDKIK 659
Qy 644 NLREVRGHLKCKRPEECCHIKISYHTQHKGRKLRH--GSSLHPRFKGLQKRD-KVMILLRE 700
Db 659 NLREVRGHLKCKRPEECSCSKQSYNKEKGVKQKELKSHLPFEKAAQEVDSKLQFKE 718
Qy 701 -OKRKKKRLKRLKRLQNDTCSMPGLTCFTHDNHOHQTAPFTWLTGPFCACTSANNTYWC 759
Db 719 NNRREKKEKKEKRGKGBECSLPLGTCFTHDNHOHQTAPFTWLTGPFCACTSANNTYWC 778
Qy 760 MXTINETHNLFCEPATGFLFYFDLNTDPYOLMNAVNTLDROVLNQLRVOLMELRSCKGY 819
Db 779 LXTVNETHNLFCCEPATGFLFYFDLNTDPYOLMNAVNTLDROVLNQLRVOLMELRSCKGY 838
Qy 820 KOCNPRTRNMDLGLKGGSYEQYQFQRRKWPMPKPPSSKSLGQWEGWEG 870
Db 839 KOCNPRPKNLDVGNKGGSYDLHR-----GQLMDGWEG 871

RESULT 4
US-09-668-673B-2
; Sequence 2, Application US/09668673B
; Patent No. 6562956
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; APPLICANT: Dhoot, Gurtej K
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PENN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 867
; TYPE: PRT
; ORGANISM: quails
US-09-668-673B-2

Query Match 65.6%; Score 3117.5; DB 4; Length 867;
Best Local Similarity 63.9%; Pred. No. 1.5e-295;
Matches 571; Conservative 117; Mismatches 146; Indels 59; Gaps 10;

Qy 5 SLVLCILSATVFSLLGGSSAFSLSHRLKGRFORDRNRIRPNILVLTDDQDVELGSMQVM 64
Db 7 ALVLAVLST---ELTSSHSSTLKSILRFRGRVQQRKIRPNILVLTDDQDVELGSLQVM 63
Qy 65 NKTRIMEOGGTHFINAFVTTMCCPSRSSILTKYVHNHNTYTNENCSPPSQAQHES 124
Db 64 NKTRIMEGSGAFINAFVTTMCCPSRSSMLTKYVHNHNTYTNENCSPPSQAQHEP 123
Qy 125 RFAVYLNSTGYRTAFPGKYLYNEYSYVPGKEMVGLLKNRSFYNTLCRNGVKEKHG 184
Db 124 RFAVYLNSTGYRTAFPGKYLYNEYSYVPGKEMVGLLKNRSFYNTLCRNGVKEKHG 183
Qy 185 SDYSKDYLTDLTNDVSVFRTSKMYPHRPVLMVISHAAPHGSDSAPQYSLRFPNASQ 244
Db 184 FQYAKDYFTDLITNESINYPKSKMYPHRPVLMVISHAAPHGSDSAPQYSLRFPNASQ 243
Qy 245 HITPSYNYAPNDKHWIMQYTGPMPIHMEFTNMLORKELOTLMSVDDSMETIYNMLVET 304
Db 244 HITPSYNYAPNDKHWIMQYTGPMPIHMEFTNMLORKELOTLMSVDDSMETIYNMLVET 303
Qy 305 GELDNTYIVYVYADHGYHIGQGLVKGSKMYPBFDIRVPYVPGVNPVAGCLNPHIVLNID 364

Db 304 GELENTYIIYTADHGVIHQFGLVKSGMPYDFDIRVPPFIRGPSVPSVVPQIVLNID 363
QY 365 LAPTIIDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKQKRVWRDPSLVERGKLKLRD 424
Db 364 LAPTIIDIAGLDTPDMDGKSLKLLDTERPGRNFRNTNKTIKWRDTPLVERGKFLRKE 423
QY 425 NDKVDAQEBNLPKYQKRVKDLQRAEYOTACBOLGKQKQCVEDATGKLKHLKHKGPMLG 484
Db 424 EANKVTQSQNLQPKYKRVKELQQAQYQACQPGKQKQCTEDASGKURIRHKCVSSDIL 483
QY 485 GSRALSNLV-PKYGGSPACTCDSDGKLSLAGRR-----KKLKKKKYKASYVRSR 537
Db 484 AIRKRTSRSHSGYSGKDKCNCGDTDFNSRTQRKNQRQFLRNPQAQKYKPRFVHTROT 543
QY 538 RSVAIEDGRVYHVLGD---AAOPRLTKKH-----WPGAPEDODDKGGDFS 583
Db 544 RLSVFEFEIYDINLEEELOVLKTRSTIKKHNAENDKKAETDGPCTMVAADGTDVI 603
QY 584 GTGGLPDYGAANPIKYTHRCYILENDTVOCDDLKYSKQAMKDKLHIDHEIETLQNKIK 643
Db 604 G-----QPSSRVYTHKCFILPNDTIRCERELQSAKAMKHKAYIDKEIALQDKIK 655
QY 644 NLREVRGHLKKRPRECDCHKISYHTQHG-----RLKHRSSSLHPPFKGLQEKD-KVWL 697
Db 656 NLREVRGHLKKRPRECDCHKISYHTQHG-----SHLHPFKBAQEVDSKLQL 712
QY 698 LREOKKKKLRKLLKRLQNNDTCSMPLGTCFTHDQHMOTAPFWTLGPFCACTSANNTY 757
Db 713 FXENRRKKERKGGKQKGGDCSLPGLTCFTHDNNHMTAPFWTLGPFCACTSSNNNTY 772
QY 758 WCMRTINETHNLFCEFATGFLFYDLNTPYQLMNAVNTLDRDVLNQLHVQLMELRSCK 817
Db 773 WCLRTVNDTHNLFCEFATGFLFYDLNTPYQLMNAVNTLDRDVLNQLHVQLMELRSCK 832
QY 818 GYKQCNPRFNDLGLKQGSYEQYRQRRKQKPEKPSKSLGOLWEGEG 870
Db 833 GYKQCNPRFNDLGLKQGSYEQYRQRRKQKPEKPSKSLGOLWEGEG 867

RESULT 5
US-09-668-673B-16
; Sequence 16, Application US/09668673B
; Patent No. 6562956
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; APPLICANT: Dhoot, Gurtej K
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; FILE REFERENCE: PERN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 1611
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (819)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (840)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (844)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (852)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (858)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (865)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (875)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (881)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (888)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (896)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (907)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (910)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (915)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (927)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (943)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (945)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (948)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (954)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (959)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (971)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE

```

; LOCATION: (974)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1018)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1046)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1080)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1089)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1102)..(1103)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1105)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1121)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1127)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1191)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1199)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1223)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1235)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1250)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1307)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1321)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1356)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1362)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1382)..(1383)

```

```

; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1397)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1431)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1437)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1448)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1458)
; OTHER INFORMATION: any amino acid
; FEATURE:

Query Match      64.4%; Score 3059; DB 4; Length 1611;
Best Local Similarity 65.6%; Pred. No. 2.1e-289;
Matches 552; Conservative 112; Mismatches 129; Indels 48; Gaps 11;

QY 55 DVELGSMQVANKTRIMEQGTHFINAFVTTMCCPSRSSILTKYVHNHNTYNNENCS 114
Db 1 DVELGSLQVANKTRIMEHGGATFINAFVTTMCCPSRSSMLTKYVHNHNTYNNENCS 60

QY 115 SPSWQAHESRTFAVYLNSTGYRTAFPGKYLNBYNGSVYPPGKEMVGLLKNSREYNTL 174
Db 61 SPSWQAHSEPRFAVYLNNTGYRTAFPGKYLNBYNGSVYPPGKREWLGLKNSRPNYTV 120

QY 175 CRNGVKEHGSYSDYKDYLTDLITNDSVSFFRTSKKYPRHPVLMVISHAAPHGSDSAPQ 234
Db 121 CRNGIKKHGFDYAKDYFTDLITNESINVFKMSKMYPRHPVMMVISHAAPHGSDSAPQ 180

QY 235 YSLFPNASCHITPSSYNYAPNDKIMRYTGPMPKPIHMEFTNMLQRLQTLMSVDSM 294
Db 181 FSKLYPNASCHITPSSYNYAPNDKIMRYTGPMPKPIHMEFTNMLQRLQTLMSVDSV 240

QY 295 ETIYNMLVETGELDNTYIVYTADHGVIHQFGLVKGKSMPEYEFDIRVPYVGRPNVEAGC 354
Db 241 ERLYNMLVETGELENTYIYTADHGVIHQFGLVKGKSMPEYDFDIRVFFPGPSVEGGS 300

QY 355 LNPHIVLINIDLAPTILDIAGLDI PADMDGKS ILKLLDTERPVNRPHLKKQMRVWRDSFLV 414
Db 301 IVPQIVLINIDLAPTILDIAGLDTPPDVDGKSVLKLLDPEKPGNRFRNTKCAKIMRDTFLV 360

QY 415 ERGKLLHKKRDNDKVDQAEENFLPKYORVKDLCORAEYOTACEQLGQKWCQVEDATGKLKL 474
Db 361 ERGKFLRKEESSKNIQSNHLPKRYVKELCQQAQRYQTACEQPGQKWCQVEDATGSKLRI 420

QY 475 HKCKGPMRLGSSR-ALSNLVPKYIYGGSSEACTCDSDYKLSLAGRR-KKLFK-----KY 527
Db 421 HKCKGPSDLLTVRQSTRNLYARGFHDKDKCESGYSRASRSQKSKORQLRQNGTPEKY 480

QY 528 KASYVRSRSIRSVAIETDGRVTHVGLD-----AAQPNLTGHWPG--APEDQDDKQGG 580
Db 481 KPRFVHTQTRSLSEFGEIYDINLEBEEELQVLPQPNIAKRHDEHGKGRDLOASSGG 540

QY 581 -----DFSGTGGPLPDYSAANPIKVTHRCYILENDTVQCOLDLYKSLQAKQKHLKHTDH 633
Db 541 NRCGMLADSSNAVGPP-----TTVRVTHKCFILPNDSTHCEBELYQSARANKHAYIDK 595

QY 634 ELETQNKIKNLRVGRHLKXKRPBECCHKISYHTQHKGRKLRH--GSSLHPRFKGQJB 691
Db 596 EIALQDKIKNLRVGRHLKXKRPBECSCSKQSYNKEKGKQKQKSHLHPFKAAQJB 655

QY 692 KD-KWLLRE-QKRRKKKRLKRLQNNDTCSPMGLTCFTHDNQHWTAPFTWLGPPFCAC 749

```

Db 656 VDSKQLFLKNNRRKKRKEKRQRKGBCSLPGLTCTFTHDNHMQTAPFNLGSFCA 715
Qy 750 TSANNNTYMCWRTTINETHNLFCEPATGFLYEYEDLNTDPYQLMNAVNTLDRDVLNQLHVQ 809
Db 716 TSSNNNTYMCWRTTINETHNLFCEPATGFLYEYEDLNTDPYQLMNAVNTLDRDVLNQLHVQ 775
Qy 810 LMELRSCGKYKQCNPRTRNMDLGLKDGGSYEQYQFQRKWPMPKRPSSKSLGQLWEGWE 869
Db 776 LMELRSCGKYKQCNPRTRNMDLGLKDGGSYEQYQFQRKWPMPKRPSSKSLGQLWEGWE 817
Qy 870 G 870
Db 818 G 818

RESULT 6
US-09-668-673B-14
; Sequence 14, Application US/09668673B
; Patent No. 6562956
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; APPLICANT: Dhoot, Gurtej K
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PENN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (142-143)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (146)
; OTHER INFORMATION: any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (445)
; OTHER INFORMATION: any amino acid
US-09-668-673B-14

Query Match 44.1%; Score 2093; DB 4; Length 455;
Best Local Similarity 86.1%; Pred. No. 8.3e-196;
Matches 384; Conservative 16; Mismatches 38; Indels 8; Gaps 4;
Qy 431 QEEFLPKYQVVKDCORAEYQACEQGLQKWCQVEDATGKLLKCKGKPMRL--GGSR 487
Db 1 QEEFLPKYQVVKDCORAEYQACEQGLQKWCQVEDATGKLLKCKGKPMRFGGGSR 60
Qy 488 ALSNLVPKYQGSSEACTCDS---GDYKLSLAGRRKLLPKGKYKASYSRVSIRSVAIEV 544
Db 61 ALSNLVPKYQGSSEACTCDSGGGGDYKLSLAGRR-KLFKKYKTSYARNRSIRSVAIEV 119
Qy 545 DGRVYHVLGDAAPRNLTKRHWPGAPEDODDQGGDFSGTGLPDSYSAAPNPKVTHRCY 604
Db 120 DGEIYHVLGDTVPQPNLSRPHXXGAXEDQDDKDGGSFSGTGLPDSYSAAPNPKVTHRCY 179
Qy 605 ILENDTVQCDLGLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKGPEECDCCHK 664
Db 180 ILENDTVQCDLGLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKGPEECDCCHK 239
Qy 665 ISYTHQKGRUKHKGSSLPFRKGLQEKDKVLLRQKRRKQKRLKLLKRLQNNDCSMPG 724
Db 240 ISYTHQKGRUKHKGSSLPFRKGLQEKDKVLLRQKRRKQKRLKLLKRLQNNDCSMPG 298

Qy 725 LTCFTHDNHMQTAPFMTLGPFCACSTANNNTYMCWRTTINETHNLFCEPATGFLYEYEDL 784
Db 299 LTCFTHDNHMQTAPLMTLGPFCACSTANNNTYMCWRTTINETHNLFCEPATGFLYEYEDL 358
Qy 785 NTDPYQLMNAVNTLDRDVLNQLHVMELRSCGKYKQCNPRTRNMDLGLKDGGSYEQYRQ 844
Db 359 STDPYQLMNAVNTLDRDVLNQLHVMELRSCGKYKQCNPRTRNMDLGLKDGGSYEQYRQ 418
Qy 845 FORKWPMPKRPSSKSLGQLWEGWE 870
Db 419 FORKWPMPKRPSSKSLGQLWEGWE 444

RESULT 7
US-09-886-319A-35
; Sequence 35, Application US/09886319A
; Patent No. 6586185
; GENERAL INFORMATION:
; APPLICANT: Wolf, Eckard
; APPLICANT: Werner, Sabine
; APPLICANT: Halle, Jorn-Peter
; APPLICANT: Regenbogen, Johannes
; APPLICANT: Goppelt, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
; TITLE OF INVENTION: Active Substances
; FILE REFERENCE: 50125/014002
; CURRENT APPLICATION NUMBER: US/09/886,319A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/222,081
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: DE 10030149.5
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 371
; OTHER INFORMATION: Xaa= any amino acid
US-09-886-319A-35

Query Match 35.8%; Score 1702; DB 4; Length 470;
Best Local Similarity 75.3%; Pred. No. 1.5e-157;
Matches 327; Conservative 21; Mismatches 36; Indels 50; Gaps 7;
Qy 300 MLVETGELDNTYIVYTADGHYHIGQGLVKGKSMYPEDIRVPFYVRGPNVYAGCLNPHI 359
Db 1 MLVETGELDNTYIVYTADGHYHIGQGLVKGKSMYPEDIRVPFYVRGPNVYAGCLNPHI 60
Qy 360 VLNIDLAP-----TILDIAGLDIPADMGGKSLKLLDTEPVN 397
Db 61 VLNIDLGPHTGYRTWGPCRHGREVYSQTTGLRAASEPVPLEKEAGL-----GETP-- 113
Qy 398 RFELKKQKRVWRDPSFLVERGKLLHKRDNDKYDAQEBNFKPYQVVKDCORAEYQACEQ 457
Db 114 -----SWMEA-----NLLHKREGDKVNAQEBNFKPYQVVKDCORAEYQACEQ 159
Qy 458 LQKWCQVEDATGKLLKCKGKPMRL---GGSRALSNLVPKYQGSSEACTCDS---GDY 511
Db 160 LQKWCQVEDATGKLLKCKGKPMRFGGGGSRLSNLVPKYQGSSEACTCDSGGGGDY 219
Qy 512 KLSLAGRRKLLPKGKYKASYSRVSIRSVAIEVDGRVYHVLGDAAPRNLTKRHWPGAP 571
Db 220 KLSLAGRR-KLFKKYKTSYARNRSIRSVAIEVDGEIYHVLGDTVPQPNLSRPHXXGAX 278
Qy 572 EDDDDKDGDFSGTGLPDSYSAAPNPKVTHRCYILENDTVQCDLGLYKSLQAWKHKLHI 631

Db 279 EDQDDKGGFSFGTGLPLDYAPNPIKVTHRCVILENTVQCOLDLYKSLQAKDKHLHI 338
QY 632 DHEIETLQNKIKNLEVRGHLKKRPECDCCHKISHTQHGRLKRCSSLHPFRKGLQE 691
Db 339 DHEIETLQNKIKNLEVRGHLKKRPECDCCHKISYHSQHKGRXKHKGSSLHPFRKGLQE 398
QY 692 KDKVLLREKQKXK 705
Db 399 KDKVLLRDRNRR 412
RESULT 8
US-09-810-347-4
; Sequence 4, Application US/09810347
; Patent No. 6461847
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01169
; CURRENT FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Human
US-09-810-347-4

Query Match 33.8%; Score 1606.5; DB 4; Length 360;
Best Local Similarity 79.8%; Pred. No. 2.2e-148; Indels 3; Gaps 1;
Matches 285; Conservative 40; Mismatches 29;
Db 5 SLVCLLSATVFSLGGSSAFSLSHRLKGRFQRRNIRPNILVLTDDQDVELGSMQVM 64
7 ALVLAVLGT---ELLGSLCSTVRSFPRGRIGQERKNIRPNILVLTDDQDVELGSLQVM 63
QY 65 NKTRIMEQGGTHFINAFVTTMCCPSSSILTKYVHNHTYNNENCSPPSQWQHES 124
Db 64 NKTRIMEHGGATFINAFVTTMCCPSSSMLTKYVHNHTYNNENCSPPSQWQHES 123
QY 125 RTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLKNRFRNYTLCRNGVKEKHG 184
Db 124 RTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLKNRFRNYTLCRNGVKEKHG 183
QY 185 SDYSKDYLTDLITNDSVSFFRTSKWYPHRPVLMVISHAAPHGPDSSAPQYSLFPNASQ 244
Db 184 FDYAKDYFTDLITNESINYPFKSKMYPHRPVLMVISHAAPHGPDSSAPQYSLFPNASQ 243
QY 245 HITPSYNYAPNPKDQKHWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 304
Db 244 HITPSYNYAPNPKDQKHWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 303
QY 305 GELDNTYIVYTDHGYHIGQGLVKGKSNPYEFDIRVPFYVGRGNVBEAGCLNPHIVL 361
Db 304 GLENTYIITADHGYHIGQGLVKGKSNPYEFDIRVPFYVGRGNVBEAGCLNPHIVL 360
US-09-810-347-2

RESULT 9
US-09-810-347-2
; Sequence 2, Application US/09810347
; Patent No. 6461847
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01169
; CURRENT FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Human
US-09-810-347-2

Query Match 33.5%; Score 1590.5; DB 4; Length 372;
Best Local Similarity 80.1%; Pred. No. 8.4e-147; Indels 3; Gaps 1;
Matches 281; Conservative 40; Mismatches 27;
QY 5 SLVCLLSATVFSLGGSSAFSLSHRLKGRFQRRNIRPNILVLTDDQDVELGSMQVM 64
Db 7 ALVLAVLGT---ELLGSLCSTVRSFPRGRIGQERKNIRPNILVLTDDQDVELGSLQVM 63
QY 65 NKTRIMEQGGTHFINAFVTTMCCPSSSILTKYVHNHTYNNENCSPPSQWQHES 124
Db 64 NKTRIMEHGGATFINAFVTTMCCPSSSMLTKYVHNHTYNNENCSPPSQWQHES 123
QY 125 RTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLKNRFRNYTLCRNGVKEKHG 184
Db 124 RTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLLKNRFRNYTLCRNGVKEKHG 183
QY 185 SDYSKDYLTDLITNDSVSFFRTSKWYPHRPVLMVISHAAPHGPDSSAPQYSLFPNASQ 244
Db 184 FDYAKDYFTDLITNESINYPFKSKMYPHRPVLMVISHAAPHGPDSSAPQYSLFPNASQ 243
QY 245 HITPSYNYAPNPKDQKHWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 304
Db 244 HITPSYNYAPNPKDQKHWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNMLVET 303
QY 305 GELDNTYIVYTDHGYHIGQGLVKGKSNPYEFDIRVPFYVGRGNVBEAGCL 355
Db 304 GLENTYIITADHGYHIGQGLVKGKSNPYEFDIRVPFYVGRGNVBEAGCL 354
US-09-668-673B-3

RESULT 10
US-09-668-673B-3
; Sequence 3, Application US/09668673B
; Patent No. 6562956
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P
; APPLICANT: Droot, Gurtej K
; TITLE OF INVENTION: IDENTIFICATION AND CLONING OF A NEW SUBFAMILY OF
; TITLE OF INVENTION: SULFATASES AND FUNCTIONAL EMBRYONIC TECHNIQUES FOR
; TITLE OF INVENTION: CHARACTERIZATION OF SUCH PROTEINS
; FILE REFERENCE: PENN-0733
; CURRENT APPLICATION NUMBER: US/09/668,673B
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,738
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 709
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-668-673B-3

Query Match 33.4%; Score 1589; DB 4; Length 709;
Best Local Similarity 40.2%; Pred. No. 3.4e-146; Indels 158; Gaps 17;
Matches 317; Conservative 118; Mismatches 196;
QY 45 NILVLTDDQDVELGSMQVMKTRRIMEQGGTHFINAFVTTMCCPSSSILTKYVHNH 104
Db 36 NVILILTDDQDIELGSLCSTVRSFPRGRIGQERKNIRPNILVLTDDQDVELGSLQVM 95
QY 105 NYTYNNENCSPPSQWQHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKWEVGLL 164
Db 96 HVHTNNQNCCTGVEWVKVHKKSIGVYLBQAGRYATYGLKYLNEYDGSYIPPGWDEWHAIV 155
QY 165 KNSRFNYTLCRNGVKEKHGSDYSKDYLTDLITNDSVSFFRTSKQYPHRPVLMVISHAA 224

US-08-484-493-11

; Sequence 11, Application US/08484493
; Patent No. 5728381
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielsicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; TITLE OF INVENTION: IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,493
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 510 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-493-11

Query Match 17.4%; Score 827.5; DB 1; Length 510;
Best Local Similarity 27.5%; Pred. No. 6.1e-72;
Matches 218; Conservative 77; Mismatches 158; Indels 341; Gaps 19;
QY 43 RNIILVLTDDDELGLSMQVMTNTRIMEQGGTHFINAFVTTMCCPSRSSILTCKYVH 102
DB 3 RNVVLLTDDDELGLSMQVMTNTRIMEQGGTHFINAFVTTMCCPSRSSILTCKYVH 62
QY 103 NHTYTN--NENCSSPSWQAQHSRTFAVYLS--TGYR--TAFEGKYLYNEYG-----S 151
DB 63 NHHVNVNTLEGNCSSKSWQIQEPNTPPAILRSGWQYQTFTFAGKYLYNEYGAPDAGGLE 122
QY 152 YVPPGKKEVGLLKSRFYNTLCRGVKEKHGSDYSKVDLTDLTNDVSP--FRTSKK 209
DB 123 HVLPLGWSYVALEKNSKTYNTLTSLNGKARKHGENYSVDYLTDLVANSVLDYLDYKSN 182
QY 210 MYPHRPVLMVISHAAPHGPDSPAPYGRPLPNASQHLTPSNVYAPNPDKHWIMRYT-GPM 268
DB 183 -----PFFMMIATPAPHSFWTAAYQYKAFQNVFAPRNKFN-IHGTNKHLLIRQAKT 236
QY 269 KPIHNEFTNMLQKRLQTLMSVDDSMETVYNNMVLVETGELDNTYIVYTADHYHIGQFGLV 328
DB 237 TNSIQFLDNAPKRKRWQTLSSVDYDLVKEVLVRLFTGELANNYVIFYTSDNGYHTGQFSLP 296

QY 329 KCKSMYPEDIRVPEYVGPVNVAGCLNPHVILNIDLAPTILDIAGLDI-PADMDGKSIL 387
DB 297 IDKRQLYEFDIKVPVLLVRGPGIKPNQTSKMLVANIDLGP-ILDIAGYDLNKTOMDGMSSL 355
QY 388 KLDDTERPVNREHLKKKVRVNRDFFLVERGKLLHKRDNDKVDAQEENFLPKYQRVKDLQC 447
DB 356 PIL---RGASNL-----TWRSVLYE----- 373
QY 448 RAEYQTACEQLGQKWQVEDATGKLLKHCKGPMRLGSSRALSNLVPKYVQGSSEACTCD 507
DB 374 -----YQEG----- 378
QY 508 SGDYKLSLAGREKLLFKKKYKASYVRSRSIRSAIEVDGRVYHVGGLGDAQAQPNMLTKRHW 567
DB 379 -----RNVT----- 382
QY 568 PGAPEDDDKOGDFSGTGGLPDYSAANPIKVTHRCYLENDTVQCOLDLYKSLQAWKH 627
DB 383 ----- 382
QY 628 KLHIDHIEITLQNKILNREVRGHLKKRPEECDCCHKISYHTQHKRGLKHGSSLSHPFRK 687
DB 383 -----DPTCPSLS----- 390
QY 688 GLQEKDKVLLREQRKKKLLKRLKRLQNNDTCSMPGLT-CFTHDNQHWQTPPTWTLGPF 746
DB 391 -----PGVSQCF-----PD 399
QY 747 CACTSANNTYWCMTINETHNLFCEF--ATGFLVEFDLNTDPVQLMNAVNTLDRDVLN 804
DB 400 CVCEDAYNTYACVTMTSALNWLQYCEFDQDEVFVENVYNTADPDQITNIAKTIDPELLG 459
QY 805 QLHVQLMELRSCKG 818
DB 460 KNNYRLMLQSCSG 473
RESULT 14
US-08-484-494-11
; Sequence 11, Application US/08484494
; Patent No. 5798239
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielsicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; TITLE OF INVENTION: IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,494
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:

```
NAME: DiGiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8416Z
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 510 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-494-11

Query Match 17.4%; Score 827.5; DB 1; Length 510;
Best Local Similarity 27.5%; Pred. No. 6.1e-72;
Matches 218; Conservative 77; Mismatches 158; Indels 341; Gaps 19;

Qy 43 RPNILVLTDDQDELGSMQVNMKTRIMEQGGTHFINAFVTPMCCPSRSSILTCKYVH 102
Db 3 RPNVLLTDDQDELGSMQVNMKTRIMEQGGTHFINAFVTPMCCPSRSSILTCKYVH 62
Qy 103 NNTYTN--NENCSSPQQAQHSRTFAVYVLS--TGVR--TAFPGKYLYNEYNG-----S 151
Db 63 NHHVNVNTLEGNCSSKQWKIOEPNTPFALLRSMQGYOTFTFFAGKYLYNEYGAPDAGGLE 122
Qy 152 YVPPGKWEVGLLKNRFRNYTLRCNGVKRKGSDYKDYLTDLITNDVSFP--FRTSKK 209
Db 123 HVPGLGWSYVALEKSKYNYTLISNGKARKHGENYSVDYLTDLVLANVSLDFLDYKSNEE 182
Qy 210 MYPHRPVLVISHAAPHGSDSAPQYSRLFPNASQHIPTPSYNAENPKHIMRYT-GPM 268
Db 183 -----PFFMMIATPAPHSMTAPQYQKAPQVAPRNNKNF--IHGTNKLWLIRQAQTPM 236
Qy 269 KPIHMEFTNKLQRKRLQTLMSVDDSMETIYNMVLVTGELDNTYIYVTADHGHIHQFGLV 328
Db 237 TNSSIQFLDQNAFRKRWOTLLSVDDLVEKLVKRLPTGELNNTYIFTSDNGYHTGQSLP 296
Qy 329 KGKSMPEYEDIRVPVVRGNEVAGCLNPHIVLNTDLAPTILDIAGLDI-PADWQGSIL 387
Db 297 IDRQYQYEPDIKVPVLLVRGPGIKPNQTSKGLVANIDLGP-ILDIAGYDLNKTQMDGMSLL 355
Qy 388 KLLDTRPVRNRLKKKMKVRSDSLFVRGKLLHKRDNKVDQAOBENFLPKYQYKOLCO 447
Db 356 PIL--RGASNL-----TWRSDDLVE----- 373
Qy 448 RASYQTACBQLGQKWQVEDATGKLLHKCKGPMRLGGSRALSNLVPKYQGSGEACTD 507
Db 374 -----YQEG----- 378
Qy 508 SGDYKLSLAGRRKKLFKKYKASYVRSRSIRSAIEVDGRVYHVLGDAAPRNLTKRHW 567
Db 379 -----RNV----- 382
Qy 568 PGAPEDQDDXGDFSGTGGLPDYSAANI KUTHRCYILENDTVQCDLILYKSLQAKWH 627
Db 383 ----- 382
Qy 628 KLHIDHIEITLQNKILNLRVGRHLKXKRPEECCHKISYHTQHKRLKRGSSLHPPK 687
Db 383 -----DPTCPSL----- 390
Qy 688 GLQEKVKVLLRSQKRRKKLKLKRLQNNDTCSMPGLT-CPTHQNHQWQAPFWTLGPF 746
Db 391 -----PGVSQCF-----PD 399
Qy 747 CACTSANNTYTCWRTINETHNLFCEP--ATGPLEYEDLNTDPYQLANVNTLDROVLN 804
Db 400 CVCEDATNNTYACVRTWSALWNLQYCEFDQEVFEVYVNTLADPDQITNIAKTIDPELLG 459
Qy 805 QLHVQLMELRSCKG 818
```

```
Db 460 KMYRLMMLQSCSG 473

RESULT 15
US-08-345-212-11
Sequence 11, Application US/08345212
Patent No. 5932211
GENERAL INFORMATION:
APPLICANT: Mileon, Peter J
APPLICANT: Morris, Charles P
APPLICANT: Anson, Donald S
APPLICANT: Occhiodoro, Teresa
APPLICANT: Bielecki, Julie
APPLICANT: Clements, Peter R
APPLICANT: Hopwood, John J
TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
TITLE OF INVENTION: IDURONATE 2-SULFATASE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/345,212
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: DiGiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8416Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 510 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-345-212-11

Query Match 17.4%; Score 827.5; DB 2; Length 510;
Best Local Similarity 27.5%; Pred. No. 6.1e-72;
Matches 218; Conservative 77; Mismatches 158; Indels 341; Gaps 19;

Qy 43 RPNILVLTDDQDELGSMQVNMKTRIMEQGGTHFINAFVTPMCCPSRSSILTCKYVH 102
Db 3 RPNVLLTDDQDELGSMQVNMKTRIMEQGGTHFINAFVTPMCCPSRSSILTCKYVH 62
Qy 103 NNTYTN--NENCSSPQQAQHSRTFAVYVLS--TGVR--TAFPGKYLYNEYNG-----S 151
Db 63 NHHVNVNTLEGNCSSKQWKIOEPNTPFALLRSMQGYOTFTFFAGKYLYNEYGAPDAGGLE 122
Qy 152 YVPPGKWEVGLLKNRFRNYTLRCNGVKRKGSDYKDYLTDLITNDVSFP--FRTSKK 209
Db 123 HVPGLGWSYVALEKSKYNYTLISNGKARKHGENYSVDYLTDLVLANVSLDFLDYKSNEE 182
Qy 210 MYPHRPVLVISHAAPHGSDSAPQYSRLFPNASQHIPTPSYNAENPKHIMRYT-GPM 268
```


Db 183 -----PFMMIATPAPHSEPWTAAPQYQKAFQNVFAPRNKNFN- IHGTNKHMLIRQAKTPM 236
QY 269 KPIHMEFTNMLQBRKLOTLSVDDSDNETIYNMLVETGELDNITYIUYTADHGYHIGQGLV 328
Db 237 TNSSIQFLDNKAFKRWQTLSSVDDLVEKLVKRLLEPTGELNNTYIIFYTSDNGYHTGQFSLP 296
QY 329 KGKSMPEFDIRVFPVVRGPNVEAGCLNPHI VINIDLAPTILDIAGLDI- PADMDGKSIL 387
Db 297 IDKRLYEFDIKVPLLVVRGPIKPNQTSKMLVANIDLGP- ILDIAGYDLNKTQMDGHSLL 355
QY 388 KLLDTERPVRNFKKKMRVVRDPSFLVERGKLLHKRDNKDVAQBEENFLPKYQVRVKLCQ 447
Db 356 PIL---RGASNL-----TWRSDDLVE----- 373
QY 448 RAEYQTACEQLGQKWQVEDATGKLLHKCKGPMRLGGSRALSNLVPKYQGSEACTCD 507
Db 374 -----YQSEG----- 378
QY 508 SGDYKLSLAGRRKKLPKKYKASYRSTRSIRSAIEVDGRVYHVGIDGAAQPRNLTGKH 567
Db 379 -----RNV----- 382
QY 568 PGAPEDQDDKGGDFSGTGLPDYSAANPIKVTHRCYILENDTVOCDDLKLYKSLQAWKDH 627
Db 383 ----- 382
QY 628 KLIHDIETLQNKIKNLEVRGHLKKGPEBCDCHKISYHTQHKGRUKHRGSSLIHPRK 687
Db 383 -----DPTCPSLS----- 390
QY 688 GLQEKDKVMLREQKRLKRLKRLQNNDTCSMPGLT-CFTHDNQHWQTAPFWTLGPF 746
Db 391 -----PGVSQCF-----PD 399
QY 747 CACTSANNTYMCRTINETHNFCEP--ATGFLYFDLNTDPYQLANAVNTLDRDVLN 804
Db 400 CVCEDAYNTYACVRTMSALWNLQYCEFDQEVFEVYVNLADPDQITNIAKTIDPELLG 459
QY 805 QLVQVLMELRSCKG 818
Db 460 KMYRLMMLQSCSG 473

Search completed: October 15, 2003, 13:10:16
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2003, 13:09:25 ; Search time 77 Seconds

(without alignments)
1820.542 Million cell updates/sec

Title: US-09-970-287-2

Perfect score: 4751

Sequence: 1 MGPPSLVLCGLSATVFLSLG.....PEMKRPSSKSLGQLWEGWEG 870

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4751	100.0	870	12	US-09-970-287-2
2	4751	100.0	885	12	US-09-970-287-12
3	4746	99.9	870	12	US-10-182-951-5
4	4746	99.9	870	12	US-10-025-966A-6
5	4746	99.9	870	12	US-10-025-966A-15
6	4746	99.9	870	12	US-10-265-071-6
7	4746	99.9	870	12	US-10-265-071-15
8	4719.5	99.3	867	9	US-09-989-722-84
9	4719.5	99.3	867	9	US-09-989-723-84
10	4719.5	99.3	867	9	US-09-989-279-84
11	4719.5	99.3	867	9	US-09-989-727-84
12	4719.5	99.3	867	10	US-09-989-731-84
13	4719.5	99.3	867	10	US-09-989-732-84
14	4719.5	99.3	867	10	US-09-991-073-84
15	4719.5	99.3	867	10	US-09-990-442-84

16	4719.5	99.3	867	10	US-09-991-163-84	Sequence 84, Appl
17	4719.5	99.3	867	10	US-09-993-604-84	Sequence 84, Appl
18	4719.5	99.3	867	10	US-09-990-456-84	Sequence 84, Appl
19	4719.5	99.3	867	10	US-09-989-721-84	Sequence 84, Appl
20	4719.5	99.3	867	10	US-09-992-598-84	Sequence 84, Appl
21	4719.5	99.3	867	10	US-09-989-293A-84	Sequence 84, Appl
22	4719.5	99.3	867	10	US-09-989-735-84	Sequence 84, Appl
23	4719.5	99.3	867	10	US-09-990-444-84	Sequence 84, Appl
24	4719.5	99.3	867	10	US-09-991-181-84	Sequence 84, Appl
25	4719.5	99.3	867	10	US-09-989-730-84	Sequence 84, Appl
26	4719.5	99.3	867	10	US-09-990-436-84	Sequence 84, Appl
27	4719.5	99.3	867	10	US-09-993-687-84	Sequence 84, Appl
28	4719.5	99.3	867	11	US-09-989-734-84	Sequence 84, Appl
29	4719.5	99.3	867	11	US-09-997-653-84	Sequence 84, Appl
30	4719.5	99.3	867	11	US-09-993-667-84	Sequence 84, Appl
31	4719.5	99.3	867	11	US-09-997-428-84	Sequence 84, Appl
32	4719.5	99.3	867	11	US-09-997-666-84	Sequence 84, Appl
33	4719.5	99.3	867	11	US-09-990-438-84	Sequence 84, Appl
34	4719.5	99.3	867	11	US-09-990-562-84	Sequence 84, Appl
35	4719.5	99.3	867	11	US-09-990-711-84	Sequence 84, Appl
36	4719.5	99.3	867	11	US-09-989-726-84	Sequence 84, Appl
37	4719.5	99.3	867	11	US-09-998-156-84	Sequence 84, Appl
38	4719.5	99.3	867	11	US-09-990-437-84	Sequence 84, Appl
39	4719.5	99.3	867	11	US-09-991-157-84	Sequence 84, Appl
40	4719.5	99.3	867	11	US-09-997-514-84	Sequence 84, Appl
41	4719.5	99.3	867	11	US-09-997-573-84	Sequence 84, Appl
42	4719.5	99.3	867	11	US-09-991-172-84	Sequence 84, Appl
43	4719.5	99.3	867	11	US-09-990-726-84	Sequence 84, Appl
44	4719.5	99.3	867	11	US-09-997-559-84	Sequence 84, Appl
45	4719.5	99.3	867	11	US-09-997-601-84	Sequence 84, Appl

ALIGNMENTS

RESULT 1
US-09-970-287-2
; Sequence 2, Application US/09970287
; Publication No. US20030166283A1
; GENERAL INFORMATION:
; APPLICANT: MEYERS, Rachel
; APPLICANT: KAPILLER-LIBERMANN, Rosana
; APPLICANT: SILOS-SANTIAGO, Inmaculada
; TITLE OF INVENTION: 22437, A NOVEL HUMAN SULFATASE AND USBS THEREFOR
; FILE REFERENCE: 10147-61U1
; CURRENT APPLICATION NUMBER: US/09/970,287
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: US 60/257,082
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-287-2

Query Match		100.0%;	Score 4751;	DB 12;	Length 870;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 870;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MGPPSLVLCGLSATVFLSLGSSAFLSHRLKGPQRDRNRIRNIIILVTDDQVELGS	60		
Db	1	MGPPSLVLCGLSATVFLSLGSSAFLSHRLKGPQRDRNRIRNIIILVTDDQVELGS	60		
Qy	61	MOVNKTREIMEQGTTHFINAFVTPMCCPSRSILTGKYVHNHTYTNNESSPSWQA	120		
Db	61	MOVNKTREIMEQGTTHFINAFVTPMCCPSRSILTGKYVHNHTYTNNESSPSWQA	120		
Qy	121	QHSRTFAVILNSTGYRTAFPGKYLNEYNGSVYPPGKWWGLLNKSRFYNYTLCRNGVK	180		
Db	121	QHSRTFAVILNSTGYRTAFPGKYLNEYNGSVYPPGKWWGLLNKSRFYNYTLCRNGVK	180		

Qy	181	EXHGS	YSDYLTDLIT	NDSVS	FFRTS	KQQT	PHRPV	LWV	ISHA	PHG	PDSAP	QYS	RLLFP	240																																										
Db	181	EXHGS	YSDYLTDLIT	NDSVS	FFRTS	KQQT	PHRPV	LWV	ISHA	PHG	PDSAP	QYS	RLLFP	240																																										
Qy	241	NASO	HITPS	YNTAP	NPKD	HWNRY	TGPKP	IHME	FTNML	QKLO	TLM	S	VDDSMET	YNM 300																																										
Db	241	NASO	HITPS	YNTAP	NPKD	HWNRY	TGPKP	IHME	FTNML	QKLO	TLM	S	VDDSMET	YNM 300																																										
Qy	301	LVET	GELDN	TYIVY	TADH	GYH	IGQ	FLVK	GSMP	YFED	IRVP	FV	RGPNV	EAC	LNP	HV 360																																								
Db	301	LVET	GELDN	TYIVY	TADH	GYH	IGQ	FLVK	GSMP	YFED	IRVP	FV	RGPNV	EAC	LNP	HV 360																																								
Qy	361	LNID	LAPT	ILDI	AGLD	IIPAD	MDG	KSIL	KLD	OTER	PVNR	FHL	KCKMR	VRDS	FLV	ERG 420																																								
Db	361	LNID	LAPT	ILDI	AGLD	IIPAD	MDG	KSIL	KLD	OTER	PVNR	FHL	KCKMR	VRDS	FLV	ERG 420																																								
Qy	421	HKRD	NK	YDA	QEN	FLPKY	QRY	VDL	IC	QRA	YQTA	CEOL	GQ	KWC	YD	B	ATG	KL	HK	CKGP 480																																				
Db	421	HKRD	NK	YDA	QEN	FLPKY	QRY	VDL	IC	QRA	YQTA	CEOL	GQ	KWC	YD	B	ATG	KL	HK	CKGP 480																																				
Qy	481	MRLG	S	RALS	NL	VPKY	YGG	S	EACT	CD	S	GDY	KLS	L	AGRR	KKL	FKCKY	K	ASY	V	RS	S	IRSV 540																																	
Db	481	MRLG	S	RALS	NL	VPKY	YGG	S	EACT	CD	S	GDY	KLS	L	AGRR	KKL	FKCKY	K	ASY	V	RS	S	IRSV 540																																	
Qy	541	A	LEV	GRV	TH	VLG	D	AA	Q	PNL	T	KH	MP	G	A	ED	ODD	X	G	G	D	F	G	T	G	G	L	P	D	Y	S	A	A	N	P	I	K	V	T 600																	
Db	541	A	LEV	GRV	TH	VLG	D	AA	Q	PNL	T	KH	MP	G	A	ED	ODD	X	G	G	D	F	G	T	G	G	L	P	D	Y	S	A	A	N	P	I	K	V	T 600																	
Qy	601	HRCY	I	LE	N	D	T	V	Q	C	D	L	D	L	Y	K	S	L	Q	A	M	K	D	H	L	I	D	H	E	I	E	T	L	O	N	K	I	X	N	L	R	E	V	R	G	H	L	K	K	K	P	E	B	C 660		
Db	601	HRCY	I	LE	N	D	T	V	Q	C	D	L	D	L	Y	K	S	L	Q	A	M	K	D	H	L	I	D	H	E	I	E	T	L	O	N	K	I	X	N	L	R	E	V	R	G	H	L	K	K	K	P	E	B	C 660		
Qy	661	D	CH	K	I	S	Y	H	T	O	H	K	R	L	K	H	G	S	S	L	H	P	R	K	L	O	B	K	O	V	N	L	L	R	E	O	K	K	K	K	L	K	L	R	L	O	N	N	D	T	C 720					
Db	661	D	CH	K	I	S	Y	H	T	O	H	K	R	L	K	H	G	S	S	L	H	P	R	K	L	O	B	K	O	V	N	L	L	R	E	O	K	K	K	K	L	K	L	R	L	O	N	N	D	T	C 720					
Qy	721	S	M	P	G	L	T	C	F	T	H	D	N	H	Q	W	T	A	P	F	T	L	G	P	F	C	A	S	T	S	A	N	N	T	Y	C	M	R	T	I	N	E	T	H	N	F	L	C	E	F	A	T	G	P	L	E 780
Db	721	S	M	P	G	L	T	C	F	T	H	D	N	H	Q	W	T	A	P	F	T	L	G	P	F	C	A	S	T	S	A	N	N	T	Y	C	M	R	T	I	N	E	T	H	N	F	L	C	E	F	A	T	G	P	L	E 780
Qy	781	Y	F	D	I	N	T	D	P	I	Q	L	M	N	A	V	N	T	L	D	R	V	N	L	H	O	L	V	M	L	S	C	K	Y	Q	C	H	R	P	R	N	M	D	L	G	K	O	G	S	S	E 840					
Db	781	Y	F	D	I	N	T</																																																	

RESULT 2

```

US-09-970-287-12
; Sequence 12, Application US/09970287
; Publication No. US20030166283A1
; GENERAL INFORMATION:
; APPLICANT: MEYERS, Rachel
; APPLICANT: GLUCKSMANN, Maria A.
; APPLICANT: KAPPELIER-LIBERMANN, Rosana
; APPLICANT: SILOS-SANTIAGO, Inmaculada
; TITLE OF INVENTION: 22437, A NOVEL HUMAN SULFATASE AND USES THEREFOR
; FILE REFERENCE: 10147-6101
; CURRENT APPLICATION NUMBER: US/09/970,287
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: US 60/257,082
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 885
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-287-12

```

Query Match	100.0%	Score	4751	DB 12	Length	885
Best Local Similarity	100.0%	Pred.	No. 0			
Matches	870	Conservative	0	Mismatches	0	Indels
Qy	1	MGPPSLVLC	LSATVFS	LGSSAFLSHHRLKGRFDRDRNRIPNII	LVLTDDQV	ELGS 60
Db	16	MGPPSLVLC	LSATVFS	LGSSAFLSHHRLKGRFDRDRNRIPNII	LVLTDDQV	ELGS 75
Qy	61	MOVNKTARR	IMEQGGTHIP	INAFVTPMCCPSRSII	ITGKYVHNHN	TYTNNCSSP
Db	76	MOVNKTARR	IMEQGGTHIP	INAFVTPMCCPSRSII	ITGKYVHNHN	TYTNNCSSP
Qy	121	QESRTFAV	ILNSTGYRTA	PPGKYLNEVNGSY	YPPGKWEVGL	LKNSRFIN
Db	136	QESRTFAV	ILNSTGYRTA	PPGKYLNEVNGSY	YPPGKWEVGL	LKNSRFIN
Qy	181	EKHGSDY	SKDYLTD	LITNDSVSFFRTSKNYPHRPVL	MIISHAAPHG	PDSAPQ
Db	196	EKHGSDY	SKDYLTD	LITNDSVSFFRTSKNYPHRPVL	MIISHAAPHG	PDSAPQ
Qy	241	NASOHIT	PSYNAVNPDK	HWIMRTGTGPMKYLHMEFTN	MLORKELO	TLMASVDS
Db	256	NASOHIT	PSYNAVNPDK	HWIMRTGTGPMKYLHMEFTN	MLORKELO	TLMASVDS
Qy	301	LVETGEL	DNTIIV	VTADHGYHIGQGLV	GKSKMPYB	FDIRVPPY
Db	316	LVETGEL	DNTIIV	VTADHGYHIGQGLV	GKSKMPYB	FDIRVPPY
Qy	361	LNIDLA	PTLID	LAGLIDP	ADMDGKSL	IKLALD
Db	376	LNIDLA	PTLID	LAGLIDP	ADMDGKSL	IKLALD
Qy	421	HKRDND	KVDAQ	BEENFLPKYQ	RVKDLQ	COAEYQ
Db	436	HKRDND	KVDAQ	BEENFLPKYQ	RVKDLQ	COAEYQ
Qy	481	MLGGS	RALS	NLVPKY	YQGS	EACTC
Db	496	MLGGS	RALS	NLVPKY	YQGS	EACTC
Qy	541	AIEVDGR	VYVH	VLGDAAP	PNRLTKRHW	PGAPEDQ
Db	556	AIEVDGR	VYVH	VLGDAAP	PNRLTKRHW	PGAPEDQ
Qy	601	HRCYT	ILEND	TVQCD	LILYKSLQ	AWKH
Db	616	HRCYT	ILEND	TVQCD	LILYKSLQ	AWKH
Qy	661	DCHKIS	YHTQ	KRGLKH	RGS	SLHPFR
Db	676	DCHKIS	YHTQ	KRGLKH	RGS	SLHPFR
Qy	721	SNPGL	TC	THDNQ	WOTAP	FWTLG
Db	736	SNPGL	TC	THDNQ	WOTAP	FWTLG
Qy	781	YFDL	MTDP	YQLMNA	VNTLDRD	VLNQLH
Db	796	YFDL	MTDP	YQLMNA	VNTLDRD	VLNQLH
Qy	841	QYRQ	FOR	KWPE	KRPP	SKSLG
Db	856	QYRQ	FOR	KWPE	KRPP	SKSLG

RESULT 3
US-10-182-951-5
; Sequence 5, Application US/10182951
; Publication No. US20030138695A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BAUGHN, Mariah R.

QY	661	DCHKISYHTQHKGLKRGSSLPFRKGLQEKDKYVWLLFEQRKKKRLKRLKRLQNDTC	720
Db	661	DCHKISYHTQHKGLKRGSSLPFRKGLQEKDKYVWLLFEQRKKKRLKRLKRLQNDTC	720
QY	721	SMRGGLTCFTHDNQHWTAPFWTLGSPCACTSANNTYWCMTTINETHNPLCEPATGPLE	780
Db	721	SMRGGLTCFTHDNQHWTAPFWTLGSPCACTSANNTYWCMTTINETHNPLCEPATGPLE	780
QY	781	YFDLNTDPYQLMNAVNTLDRDVLNLQHLVQLMELRSCKYKQCNPRTRNNDLGLDKGGSYE	840
Db	781	YFDLNTDPYQLMNAVNTLDRDVLNLQHLVQLMELRSCKYKQCNPRTRNNDLGLDKGGSYE	840
QY	841	QYQFQRKQKPEMKRPSKSLGQLWEGWEG	870
Db	841	QYQFQRKQKPEMKRPSKSLGQLWEGWEG	870
RESULT 4			
US-10-025-966A-6			
; Sequence 6, Application US/10025966A			
; Publication No. US20030148920A1			
; GENERAL INFORMATION:			
; APPLICANT: Rosen, Steven D.			
; APPLICANT: Palmeri, Diana			
; APPLICANT: Stefan, Hemmerich			
; TITLE OF INVENTION: Sulfatases and methods of use			
; TITLE OF INVENTION: Sulfatases and methods of use			
; FILE REFERENCE: UCAL230			
; CURRENT APPLICATION NUMBER: US/10/025,966A			
; CURRENT FILING DATE: 2002-12-21			
; PRIOR APPLICATION NUMBER: 60/258,577			
; PRIOR FILING DATE: 2000-12-27			
; PRIOR APPLICATION NUMBER: 60/267,831			
; PRIOR FILING DATE: 2001-09-02			
; NUMBER OF SEQ ID NOS: 26			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 6			
; LENGTH: 870			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-025-966A-6			
Query Match 99.9%; Score 4746; DB 12; Length 870;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	MGPPSLVCLLSATVFSLLGSSAFLSHRLKGRFDRNRIRPNIILVLTDQDQVELGS	60
Db	1	MGPPSLVCLLSATVFSLLGSSAFLSHRLKGRFDRNRIRPNIILVLTDQDQVELGS	60
QY	61	MQVNNKTRIMEQGHFIINAFVTTMPCSPSSILTGKYYVNNHTYTNNECSPSQQA	120
Db	61	MQVNNKTRIMEQGHFIINAFVTTMPCSPSSILTGKYYVNNHTYTNNECSPSQQA	120
QY	121	QHSRTFAVILNSTGYRTAFPGKYLNEYGSYVPPCKWEKGVLLKNSRFYNYTLCRNGVK	180
Db	121	QHSRTFAVILNSTGYRTAFPGKYLNEYGSYVPPCKWEKGVLLKNSRFYNYTLCRNGVK	180
QY	181	EKHGSDYSKDYLTDLITNDSVSFFRTSKOYPHRPVLAVISHAAHPGEDSNAPQYSLRFP	240
Db	181	EKHGSDYSKDYLTDLITNDSVSFFRTSKOYPHRPVLAVISHAAHPGEDSNAPQYSLRFP	240
QY	241	NASCHITPSYNAFNPDKIMRYTGTPMKPIHMEFTNMLQRKLOTLMASVDDSMETIYNN	300
Db	241	NASCHITPSYNAFNPDKIMRYTGTPMKPIHMEFTNMLQRKLOTLMASVDDSMETIYNN	300
QY	301	LVETGELDNTYIVVTADHGYHIQFGLVKGSKSPYEPDIRVPYVRGNVEAGCLNPHIV	360
Db	301	LVETGELDNTYIVVTADHGYHIQFGLVKGSKSPYEPDIRVPYVRGNVEAGCLNPHIV	360
QY	361	LNIDLAPTILDIAGLDIPADMCKGSKILKLDLTERPVNRFHLKQGVVRDPSFLVERGKLL	420

```

Db 361 LNTDLAPTLDIAGLDIPADMDGKSLKLLKLDTERPVRPHLKKORVWRDSFLVERGKLL 420
Qy 421 HKRDNDKVDQAQENFLPKYQRYVKDLCQRAEYQTAQCEQLGQKQWQCVEDATGKLLHKCKGP 480
Db 421 HKRDNDKVDQAQENFLPKYQRYVKDLCQRAEYQTAQCEQLGQKQWQCVEDATGKLLHKCKGP 480
Qy 481 MRLGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKLLPKKKYKASYVRSRSIRSV 540
Db 481 MRLGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKLLPKKKYKASYVRSRSIRSV 540
Qy 541 AIEVDGRVHVGLGDAAPRNLTKEHWPFGAPEDQDDKQGGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVHVGLGDAAPRNLTKEHWPFGAPEDQDDKQGGDFSGTGGLPDYSAANPIKVT 600
Qy 601 HRCYILENDTVQCDLJLYKSLQAKHKLHIDHETLQNKIKMLREVVRGHLKKRPEEC 660
Db 601 HRCYILENDTVQCDLJLYKSLQAKHKLHIDHETLQNKIKMLREVVRGHLKKRPEEC 660
Qy 661 DCHKISYHTQHKGRLLKRGSSLLHPFRKGLQKQKVMLLREQKRRKLLKLLKRLQNNDT 720
Db 661 DCHKISYHTQHKGRLLKRGSSLLHPFRKGLQKQKVMLLREQKRRKLLKLLKRLQNNDT 720
Qy 721 SMPGLTCFTHDNQHMOTAPFTWLTGSPFCACTSANNNTYWCMTINETHNFLFCFATGFL 780
Db 721 SMPGLTCFTHDNQHMOTAPFTWLTGSPFCACTSANNNTYWCMTINETHNFLFCFATGFL 780
Qy 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGKYQKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGKYQKQCNPRTRNMDLGLKDGGSYE 840
Qy 841 QYRQFORRWKPEMKRPSKSLGQLWEGWEG 870
Db 841 QYRQFORRWKPEMKRPSKSLGQLWEGWEG 870

RESULT 5
US-10-025-966A-15
; Sequence 15, Application US/10025966A
; Publication No. US20030148920A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Palmeri, Diana
; APPLICANT: Stefan, Hemmerich
; TITLE OF INVENTION: Sulfatases and methods of use
; TITLE OF INVENTION: thereof
; FILE REFERENCE: UCAL230
; CURRENT APPLICATION NUMBER: US/10/025,966A
; CURRENT FILING DATE: 2002-12-21
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-966A-15

Query Match 99.9%; Score 4746; DB 12; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPPSLVLCLSATVPSLGGSSAFSLSHRLKGRFQDRDRIRNIIVLVTDQDVELGS 60
Db 1 MGPPSLVLCLSATVPSLGGSSAFSLSHRLKGRFQDRDRIRNIIVLVTDQDVELGS 60
Qy 61 MQVMNTRIMEGGGTHFINAVPTTMCPCPSRSSLTGKYVHNENTYTNENCSQSWQA 120
Db 61 MQVMNTRIMEGGGTHFINAVPTTMCPCPSRSSLTGKYVHNENTYTNENCSQSWQA 120
Qy 121 QHSRTFAVYLNSTGYRTAPFGKYLNEYNGSVYPPGKEMVGLLKNRFRNYTLRCRNGVK 180

```

```

Db 121 QHSRTFAVYLNSTGYRTAPFGKYLNEYNGSVYPPGKEMVGLLKNRFRNYTLRCRNGVK 180
Qy 181 EKSGSDYSKYLTLTNDSDVSFFRTSKQMYPHRPVLMVISHAAPHGPDSPAQYSLRFP 240
Db 181 EKSGSDYSKYLTLTNDSDVSFFRTSKQMYPHRPVLMVISHAAPHGPDSPAQYSLRFP 240
Qy 241 NASQHTTPSYNTAPNPDGKINIMRYTGPMPKPIMHEFTNMLQRRLQTLMSVDDSHETTINM 300
Db 241 NASQHTTPSYNTAPNPDGKINIMRYTGPMPKPIMHEFTNMLQRRLQTLMSVDDSHETTINM 300
Qy 301 LVETGELDNTYIVYTADHGTHIGQFGLVKGKSMPEYFDIRVPFYVRGPNVZEAGCLNPHIV 360
Db 301 LVETGELDNTYIVYTADHGTHIGQFGLVKGKSMPEYFDIRVPFYVRGPNVZEAGCLNPHIV 360
Qy 361 LNTDLAPTLDIAGLDIPADMDGKSLKLLKLDTERPVRPHLKKORVWRDSFLVERGKLL 420
Db 361 LNTDLAPTLDIAGLDIPADMDGKSLKLLKLDTERPVRPHLKKORVWRDSFLVERGKLL 420
Qy 421 HKRDNDKVDQAQENFLPKYQRYVKDLCQRAEYQTAQCEQLGQKQWQCVEDATGKLLHKCKGP 480
Db 421 HKRDNDKVDQAQENFLPKYQRYVKDLCQRAEYQTAQCEQLGQKQWQCVEDATGKLLHKCKGP 480
Qy 481 MRLGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKLLPKKKYKASYVRSRSIRSV 540
Db 481 MRLGGSRALSNLVPKYGGSEACTCDSGDYKLSLAGRRKLLPKKKYKASYVRSRSIRSV 540
Qy 541 AIEVDGRVHVGLGDAAPRNLTKEHWPFGAPEDQDDKQGGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVHVGLGDAAPRNLTKEHWPFGAPEDQDDKQGGDFSGTGGLPDYSAANPIKVT 600
Qy 601 HRCYILENDTVQCDLJLYKSLQAKHKLHIDHETLQNKIKMLREVVRGHLKKRPEEC 660
Db 601 HRCYILENDTVQCDLJLYKSLQAKHKLHIDHETLQNKIKMLREVVRGHLKKRPEEC 660
Qy 661 DCHKISYHTQHKGRLLKRGSSLLHPFRKGLQKQKVMLLREQKRRKLLKLLKRLQNNDT 720
Db 661 DCHKISYHTQHKGRLLKRGSSLLHPFRKGLQKQKVMLLREQKRRKLLKLLKRLQNNDT 720
Qy 721 SMPGLTCFTHDNQHMOTAPFTWLTGSPFCACTSANNNTYWCMTINETHNFLFCFATGFL 780
Db 721 SMPGLTCFTHDNQHMOTAPFTWLTGSPFCACTSANNNTYWCMTINETHNFLFCFATGFL 780
Qy 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGKYQKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGKYQKQCNPRTRNMDLGLKDGGSYE 840
Qy 841 QYRQFORRWKPEMKRPSKSLGQLWEGWEG 870
Db 841 QYRQFORRWKPEMKRPSKSLGQLWEGWEG 870

RESULT 6
US-10-265-071-6
; Sequence 6, Application US/10265071
; Publication No. US20030147875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; APPLICANT: Tomita, Megumi
; TITLE OF INVENTION: Sulfatases and methods of use thereof
; FILE REFERENCE: UCAL-230CIP
; CURRENT APPLICATION NUMBER: US/10/265,071
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; PRIOR APPLICATION NUMBER: 10/025,966
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6

```

```

; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-265-071-6

Query Match      99.9%; Score 4746; DB 12; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPPSLVCLLSATVPSLLGSSAFSLSHRLKGRFQDRNRIRPNILVLTDDQDVELGS 60
Db 1 MGPPSLVCLLSATVPSLLGSSAFSLSHRLKGRFQDRNRIRPNILVLTDDQDVELGS 60
Qy 61 MQVMNTRIMEGGTHFINAFVTPMCCPSRSSILTKGYVHNHTYTNNECSPSQA 120
Db 61 MQVMNTRIMEGGTHFINAFVTPMCCPSRSSILTKGYVHNHTYTNNECSPSQA 120
Qy 121 QHESRTFAVYLNSTGRTAFPGKYLNEYNGSVYPPGKEMWGLKNSRFNYTLCRNGVK 180
Db 121 QHESRTFAVYLNSTGRTAFPGKYLNEYNGSVYPPGKEMWGLKNSRFNYTLCRNGVK 180
Qy 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGDESAPOYSRLFP 240
Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGDESAPOYSRLFP 240
Qy 241 NASOHITPSTNYAPNPKHWMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db 241 NASOHITPSTNYAPNPKHWMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Qy 301 LVETGELDNITYIVYTADHGHYHIGQFGLVKGSMPEYEDIRVPFYVGRGNVEAGCLNPHIV 360
Db 301 LVETGELDNITYIVYTADHGHYHIGQFGLVKGSMPEYEDIRVPFYVGRGNVEAGCLNPHIV 360
Qy 361 LNIDLAPTILDIAGLDIPADMKGSKILKLLDTERPVRNPHLKKKRVWRDSFLVERGKLL 420
Db 361 LNIDLAPTILDIAGLDIPADMKGSKILKLLDTERPVRNPHLKKKRVWRDSFLVERGKLL 420
Qy 421 HKRDNDKVDAAQENFLPKYQVNDLCORAEYOTACEQLGQKWQCVEDATGKLHKCKGP 480
Db 421 HKRDNDKVDAAQENFLPKYQVNDLCORAEYOTACEQLGQKWQCVEDATGKLHKCKGP 480
Qy 481 MRGGSRALSNLVPKYGGSEACTCDSDGYKLSLAGRRKKLKKYKASVRSRSIRSV 540
Db 481 MRGGSRALSNLVPKYGGSEACTCDSDGYKLSLAGRRKKLKKYKASVRSRSIRSV 540
Qy 541 AIEVDGRVYHVLGDAQAQPRNLTKRHWPGAPEDQDDKDGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVYHVLGDAQAQPRNLTKRHWPGAPEDQDDKDGDFSGTGGLPDYSAANPIKVT 600
Qy 601 HRCYIILENDTVQCDDLDLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKRPEEC 660
Db 601 HRCYIILENDTVQCDDLDLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKRPEEC 660
Qy 721 SMPGLTCFTHDNQHWOTAPFWTLGPPCACTSANNTYWCMTINETHNPLFCFATGPLE 780
Db 721 SMPGLTCFTHDNQHWOTAPFWTLGPPCACTSANNTYWCMTINETHNPLFCFATGPLE 780
Qy 781 YFDLNTDPYQLMNAVNTLDRDLNQLHVLQVJMELRCKGYKQCNPRTRNDLGLKDGGSYE 840
Db 781 YFDLNTDPYQLMNAVNTLDRDLNQLHVLQVJMELRCKGYKQCNPRTRNDLGLKDGGSYE 840
Qy 841 QYRQFORRWKPEMKRPSKSLGQLMWEGWEG 870
Db 841 QYRQFORRWKPEMKRPSKSLGQLMWEGWEG 870

```

RESULT 7
JS-10-265-071-15
, Sequence 15, Application US/10265071

```

; Publication NO. US20030147875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; APPLICANT: Tomita, Megumi
; TITLE OF INVENTION: Sulfatases and methods of use thereof
; FILE REFERENCE: UCAL-230CIP
; CURRENT APPLICATION NUMBER: US/10/265,071
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; PRIOR APPLICATION NUMBER: 10/025,966
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 870
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-265-071-15

Query Match      99.9%; Score 4746; DB 12; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPPSLVCLLSATVPSLLGSSAFSLSHRLKGRFQDRNRIRPNILVLTDDQDVELGS 60
Db 1 MGPPSLVCLLSATVPSLLGSSAFSLSHRLKGRFQDRNRIRPNILVLTDDQDVELGS 60
Qy 61 MQVMNTRIMEGGTHFINAFVTPMCCPSRSSILTKGYVHNHTYTNNECSPSQA 120
Db 61 MQVMNTRIMEGGTHFINAFVTPMCCPSRSSILTKGYVHNHTYTNNECSPSQA 120
Qy 121 QHESRTFAVYLNSTGRTAFPGKYLNEYNGSVYPPGKEMWGLKNSRFNYTLCRNGVK 180
Db 121 QHESRTFAVYLNSTGRTAFPGKYLNEYNGSVYPPGKEMWGLKNSRFNYTLCRNGVK 180
Qy 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGDESAPOYSRLFP 240
Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGDESAPOYSRLFP 240
Qy 241 NASOHITPSTNYAPNPKHWMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Db 241 NASOHITPSTNYAPNPKHWMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
Qy 301 LVETGELDNITYIVYTADHGHYHIGQFGLVKGSMPEYEDIRVPFYVGRGNVEAGCLNPHIV 360
Db 301 LVETGELDNITYIVYTADHGHYHIGQFGLVKGSMPEYEDIRVPFYVGRGNVEAGCLNPHIV 360
Qy 361 LNIDLAPTILDIAGLDIPADMKGSKILKLLDTERPVRNPHLKKKRVWRDSFLVERGKLL 420
Db 361 LNIDLAPTILDIAGLDIPADMKGSKILKLLDTERPVRNPHLKKKRVWRDSFLVERGKLL 420
Qy 421 HKRDNDKVDAAQENFLPKYQVNDLCORAEYOTACEQLGQKWQCVEDATGKLHKCKGP 480
Db 421 HKRDNDKVDAAQENFLPKYQVNDLCORAEYOTACEQLGQKWQCVEDATGKLHKCKGP 480
Qy 481 MRGGSRALSNLVPKYGGSEACTCDSDGYKLSLAGRRKKLKKYKASVRSRSIRSV 540
Db 481 MRGGSRALSNLVPKYGGSEACTCDSDGYKLSLAGRRKKLKKYKASVRSRSIRSV 540
Qy 541 AIEVDGRVYHVLGDAQAQPRNLTKRHWPGAPEDQDDKDGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVYHVLGDAQAQPRNLTKRHWPGAPEDQDDKDGDFSGTGGLPDYSAANPIKVT 600
Qy 601 HRCYIILENDTVQCDDLDLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKRPEEC 660
Db 601 HRCYIILENDTVQCDDLDLYKSLQAWKHKLHIDHEIETLQNKIKNLRVGRHLKKRPEEC 660
Qy 661 DCHKISYHTQHKGRKLRGSSSLHFPKRLQEKQKQVWLLREOKRKKLKLRLQNNDTC 720
Db 661 DCHKISYHTQHKGRKLRGSSSLHFPKRLQEKQKQVWLLREOKRKKLKLRLQNNDTC 720
Qy 721 SMPGLTCFTHDNQHWOTAPFWTLGPPCACTSANNTYWCMTINETHNPLFCFATGPLE 780
Db 721 SMPGLTCFTHDNQHWOTAPFWTLGPPCACTSANNTYWCMTINETHNPLFCFATGPLE 780
Qy 781 YFDLNTDPYQLMNAVNTLDRDLNQLHVLQVJMELRCKGYKQCNPRTRNDLGLKDGGSYE 840
Db 781 YFDLNTDPYQLMNAVNTLDRDLNQLHVLQVJMELRCKGYKQCNPRTRNDLGLKDGGSYE 840
Qy 841 QYRQFORRWKPEMKRPSKSLGQLMWEGWEG 870
Db 841 QYRQFORRWKPEMKRPSKSLGQLMWEGWEG 870

```

Db 661 DCHKISYHTQHKGLKRGSSLPFRKGLQEKDKVLLRQKQKKLKLKRLQNNDIC 720
QY 721 SMGLGCTFDHNDHWTAPFTWLGPFCACTSANNTYWCRTINETHNPLCFCEATGELE 780
Db 721 SMGLGCTFDHNDHWTAPFTWLGPFCACTSANNTYWCRTINETHNPLCFCEATGELE 780
QY 781 YFDLNTDPYQLMAVNTLDRDVLNQLRVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDPYQLMAVNTLDRDVLNQLRVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840
QY 841 QYQFORRKWPENKRPSSKSLGOLWEGWG 870
Db 841 QYQFORRKWPENKRPSSKSLGOLWEGWG 870

RESULT 8

US-09-989-722-84

; Sequence 84, Application US/09989722

; Patent No. US20020072067A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Deenoyers, Luc

; APPLICANT: Eston, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kijavlin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: P2730P1063

; CURRENT APPLICATION NUMBER: US/09/989, 722

; CURRENT FILING DATE: 2001-11-19

; PRIOR APPLICATION NUMBER: 60/049787

; PRIOR FILING DATE: 1997-06-16

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/065186

; PRIOR FILING DATE: 1997-11-12

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066770

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/075945

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: 60/078910

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28

; PRIOR APPLICATION NUMBER: 60/084600

; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/087106

; PRIOR FILING DATE: 1998-05-28

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087609

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/089598

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089599

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089600

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089653

; PRIOR FILING DATE: 1998-06-17

; PRIOR APPLICATION NUMBER: 60/089801

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/089907

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/089908

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908

; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089947
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089948
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089952
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/090246
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090252
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090254
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090349
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090355
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090429
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090431
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090435
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090444
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090472
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090535
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090540
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090542
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090676
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090678
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090690
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090862
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092192
 ; PRIOR FILING DATE: 1998-07-09

Query Match 99.3%; Score 4719.5; DB 9; Length 867;

Best Local Similarity 99.5%; Pred. No. 0;
 Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
 QY 1 MGPPSLVLCILSATVPSLLGGSSAFSLSHRLKGRFORRRNIRPNIIILVLTDDQDVEIGS 60
 DB 1 MGPPSLVLCILSATVPSLLGGSSAFSLSHRLKGRFORRRNIRPNIIILVLTDDQDVEIGS 60
 QY 61 MQVMNKTRIRIMEQGGTHFINAFVITPMCCPSRSSILTKYVHNHTYTNNECSPSQA 120
 DB 61 MQVMNKTRIRIMEQGGTHFINAFVITPMCCPSRSSILTKYVHNHTYTNNECSPSQA 120
 QY 121 QHESRTFAVILNSTGTRTAPFOKYLNEYNGSVPPGKWEVGLLKNRFPNTYTCRNGVK 180
 DB 121 QHESRTFAVILNSTGYRTAFPGKYLNEYNGSVPPGKWEVGLLKNRFPNTYTCRNGVK 180
 QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRVLAVISHAAPHGPEDESAPOVSRLEP 240
 DB 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRVLAVISHAAPHGPEDESAPOVSRLEP 240
 QY 241 NASOHITPSINYAPNDPKHMIMRYTGPMPKPIHMEFTNNLORKRLQTLMSVDDSMETIYNN 300
 DB 241 NASOHITPSINYAPNDPKHMIMRYTGPMPKPIHMEFTNNLORKRLQTLMSVDDSMETIYNN 300
 QY 301 LYETGELDNITYIVYTADHGYTHIGQFGLVKGSMPEYFDIRVPFYVYRGPNVEAGCLNPHIV 360
 DB 301 LYETGELDNITYIVYTADHGYTHIGQFGLVKGSMPEYFDIRVPFYVYRGPNVEAGCLNPHIV 360
 QY 361 LNITDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKCKMRVWRDSFLVERGKLL 420
 DB 361 LNITDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRFHLKCKMRVWRDSFLVERGKLL 420
 QY 421 HKRDNDKVDQAQENFLPKYQYRVKDLICORAEYOTACEOLGOKWOCVEDATGKLKHKCKGP 480
 DB 421 HKRDNDKVDQAQENFLPKYQYRVKDLICORAEYOTACEOLGOKWOCVEDATGKLKHKCKGP 480
 QY 481 MRLGGSRALSNLVPKYVYGGSEACTCDSDGYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540
 DB 481 MRLGGSRALSNLVPKYVYGGSEACTCDSDGYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540
 QY 541 ALEVDRVYHVLGDAAPRNLTKRHWFGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
 DB 541 ALEVDRVYHVLGDAAPRNLTKRHWFGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
 QY 601 HRCYILENDTVQCDLDLYKSLQAKDHLIDHEITLQNKILNLEVRGHLKKRPEEC 660
 DB 601 HRCYILENDTVQCDLDLYKSLQAKDHLIDHEITLQNKILNLEVRGHLKKRPEEC 660
 QY 661 DCHKISYHTQHKGRKLRGSSLHPFRKGLQEKQKVMLLREOQKRLKRLQNNDDTC 720
 DB 661 DCHKISYHTQHKGRKLRGSSLHPFRKGLQEKQKVMLLREOQKRLKRLQNNDDTC 720
 QY 721 SMPGLTCFTHDQHWOTAPFWTLGPFCACTSANNNYTCMETINETHNPLFCFATGFL 780
 DB 721 SMPGLTCFTHDQHWOTAPFWTLGPFCACTSANNNYTCMETINETHNPLFCFATGFL 780
 QY 781 YFDLNTDYPQLMNAVNTLDRVNLQHLVQLMELRSCKYKQCNPRTRNMDLGLKDGSSYE 840
 DB 781 YFDLNTDYPQLMNAVNTLDRVNLQHLVQLMELRSCKYKQCNPRTRNMDLGLKDGSSYE 840
 QY 841 QYRQFORRWKPEMKRPSKSLQQLWEGWEG 870
 DB 838 QYRQFORRWKPEMKRPSKSLQQLWEGWEG 867

RESULT 9
 US-09-989-723-84
 ; Sequence 84, Application US/09989723
 ; Patent No. US20020072092A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc

```

; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C62
; CURRENT APPLICATION NUMBER: US/09/989,723
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24

```

```

; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match          99.3%; Score 4719.5; DB 9; Length 867;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 MGPPSLVLCILSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIIILVLTDDQDVELGS 60
DB 1 MGPPSLVLCILSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNIIILVLTDDQDVELGS 60
QY 61 MQWNNKTRIMEQGGTHINAFVITPMPCCPSRSSILTKGYVHNHTYTNNECSPSQWA 120
DB 61 MQWNNKTRIMEQGAHINAFVITPMPCCPSRSSILTKGYVHNHTYTNNECSPSQWA 120
QY 121 QHESFTFAYLNTGYRTAFPGKYLNBYNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 180
DB 121 QHESFTFAYLNTGYRTAFPGKYLNBYNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 180
QY 181 EKHGSDYKDYLTDLITDNDVSFRTSKMYPHRPLVLMVISHAAPHGPDSPAPQSLFP 240
DB 181 EKHGSDYKDYLTDLITDNDVSFFRTSKMYPHRPLVLMVISHAAPHGPDSPAPQSLFP 240
QY 241 NASQHTTFSYNAPNPKDHWIMRYTGPMPKPIHMEFTNNLQKRLQTLMSVDDSMETIYNM 300
DB 241 NASQHTTFSYNAPNPKDHWIMRYTGPMPKPIHMEFTNNLQKRLQTLMSVDDSMETIYNM 300
QY 301 LVETGELONTYIVYTADHGYHIGQGLVKGKSMPEYEDIRVPFYVRGNVEAGCLNPHIV 360
```

```

DB 301 LVETGELONTYIVYTADHGYHIGQGLVKGKSMPEYEDIRVPFYVRGNVEAGCLNPHIV 360
QY 361 LNIIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVRNRFHLKKQQRVWRDSFLVERGKLL 420
DB 361 LNIIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVRNRFHLKKQQRVWRDSFLVERGKLL 420
QY 421 HKRDNKYDAQSENFPLPKYQRYKOLCQRAEYQTAQCEQLGQKMQCQVEDATGKLKHLCKGKP 480
DB 421 HKRDNKYDAQSENFPLPKYQRYKOLCQRAEYQTAQCEQLGQKMQCQVEDATGKLKHLCKGKP 480
QY 481 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKLLFKKKYKASYVRSRSRSV 540
DB 481 MRLGGSRALSNLVPKYQGGSEACTCDSGDYKLSLAGRRKLLFKKKYKASYVRSRSRSV 540
QY 541 AIEVDGRVYHVLGDAAQPRNLTKRHWFGAPEDQDDKGGDPSTGGGLPDYSAANPIKVT 600
DB 541 AIEVDGRVYHVLGDAAQPRNLTKRHWFGAPEDQDDKGGDPSTGGGLPDYSAANPIKVT 600
QY 601 HRCYILENDTVQCDDLDLYKSLQANKDHKLHIDHETLQNKIKNLREVRGHLKKRPREEC 660
DB 601 HRCYILENDTVQCDDLDLYKSLQANKDHKLHIDHETLQNKIKNLREVRGHLKKRPREEC 660
QY 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQEKQKVLLEQKRRKKLKLRLQNNDTG 720
DB 661 DCHKISYHTQHKGRLLKRGSSLHPFRKGLQEKQKVLLEQKRRKKLKLRLQNNDTG 720
QY 721 SMPGLTCTHNDNQHWQTAFFWTILGPFCACTSANNTYWCMTINETHNLFCEPATGFE 780
DB 721 SMPGLTCTHNDNQHWQTAFFWTILGPFCACTSANNTYWCMTINETHNLFCEPATGFE 780
QY 781 YFDLNTDPYQLMNAVTILDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840
DB 781 YFDLNTDPYQLMNAVTILDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 837
QY 841 QYRQFORRWKPEMKRPPSSKSLGQLWEGWEG 870
DB 838 QYRQFORRWKPEMKRPPSSKSLGQLWEGWEG 867

RESULT 10
US-09-989-279-84
; Sequence 84, Application US/09989279
; Patent No. US20020072496A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC56
; CURRENT APPLICATION NUMBER: US/09/989,279
```


[illegible]

1 PRIOR FILING DATE: 1998-06-02
2 PRIOR APPLICATION NUMBER: 60/087759
3 PRIOR FILING DATE: 1998-06-02
4 PRIOR APPLICATION NUMBER: 60/087827
5 PRIOR FILING DATE: 1998-06-03
6 PRIOR APPLICATION NUMBER: 60/088021
7 PRIOR FILING DATE: 1998-06-04
8 PRIOR APPLICATION NUMBER: 60/088025
9 PRIOR FILING DATE: 1998-06-04
10 PRIOR APPLICATION NUMBER: 60/088026
11 PRIOR FILING DATE: 1998-06-04
12 PRIOR APPLICATION NUMBER: 60/088028
13 PRIOR FILING DATE: 1998-06-04
14 PRIOR APPLICATION NUMBER: 60/088029
15 PRIOR FILING DATE: 1998-06-04
16 PRIOR APPLICATION NUMBER: 60/088030
17 PRIOR FILING DATE: 1998-06-04
18 PRIOR APPLICATION NUMBER: 60/088033
19 PRIOR FILING DATE: 1998-06-04
20 PRIOR APPLICATION NUMBER: 60/088326
21 PRIOR FILING DATE: 1998-06-04
22 PRIOR APPLICATION NUMBER: 60/088167
23 PRIOR FILING DATE: 1998-06-05
24 PRIOR APPLICATION NUMBER: 60/088202
25 PRIOR FILING DATE: 1998-06-05
26 PRIOR APPLICATION NUMBER: 60/088212
27 PRIOR FILING DATE: 1998-06-05
28 PRIOR APPLICATION NUMBER: 60/088217
29 PRIOR FILING DATE: 1998-06-05
30 PRIOR APPLICATION NUMBER: 60/088655
31 PRIOR FILING DATE: 1998-06-09
32 PRIOR APPLICATION NUMBER: 60/088734
33 PRIOR FILING DATE: 1998-06-10
34 PRIOR APPLICATION NUMBER: 60/086738
35 PRIOR FILING DATE: 1998-06-10
36 PRIOR APPLICATION NUMBER: 60/088742
37 PRIOR FILING DATE: 1998-06-10
38 PRIOR APPLICATION NUMBER: 60/086810
39 PRIOR FILING DATE: 1998-06-10
40 PRIOR APPLICATION NUMBER: 60/088824
41 PRIOR FILING DATE: 1998-06-10
42 PRIOR APPLICATION NUMBER: 60/088826
43 PRIOR FILING DATE: 1998-06-10
44 PRIOR APPLICATION NUMBER: 60/088858
45 PRIOR FILING DATE: 1998-06-11
46 PRIOR APPLICATION NUMBER: 60/088861
47 PRIOR FILING DATE: 1998-06-11
48 PRIOR APPLICATION NUMBER: 60/088876
49 PRIOR FILING DATE: 1998-06-11
50 PRIOR APPLICATION NUMBER: 60/089105
51 PRIOR FILING DATE: 1998-06-12
52 PRIOR APPLICATION NUMBER: 60/089440
53 PRIOR FILING DATE: 1998-06-16
54 PRIOR APPLICATION NUMBER: 60/089512
55 PRIOR FILING DATE: 1998-06-16
56 PRIOR APPLICATION NUMBER: 60/089514
57 PRIOR FILING DATE: 1998-06-16
58 PRIOR APPLICATION NUMBER: 60/089532
59 PRIOR FILING DATE: 1998-06-17
60 PRIOR APPLICATION NUMBER: 60/089538
61 PRIOR FILING DATE: 1998-06-17
62 PRIOR APPLICATION NUMBER: 60/089598
63 PRIOR FILING DATE: 1998-06-17
64 PRIOR APPLICATION NUMBER: 60/089599
65 PRIOR FILING DATE: 1998-06-17
66 PRIOR APPLICATION NUMBER: 60/089600
67 PRIOR FILING DATE: 1998-06-17
68 PRIOR APPLICATION NUMBER: 60/089653
69 PRIOR FILING DATE: 1998-06-17
70 PRIOR APPLICATION NUMBER: 60/089801
71 PRIOR FILING DATE: 1998-06-18
72 PRIOR APPLICATION NUMBER: 60/089907
73 PRIOR FILING DATE: 1998-06-18
74 PRIOR APPLICATION NUMBER: 60/089908
75 PRIOR FILING DATE: 1998-06-18
76 PRIOR APPLICATION NUMBER: 60/089947
77 PRIOR FILING DATE: 1998-06-19
78 PRIOR APPLICATION NUMBER: 60/089948
79 PRIOR FILING DATE: 1998-06-19
80 PRIOR APPLICATION NUMBER: 60/089952
81 PRIOR FILING DATE: 1998-06-19
82 PRIOR APPLICATION NUMBER: 60/090246
83 PRIOR FILING DATE: 1998-06-22
84 PRIOR APPLICATION NUMBER: 60/090252
85 PRIOR FILING DATE: 1998-06-22
86 PRIOR APPLICATION NUMBER: 60/090254
87 PRIOR FILING DATE: 1998-06-22
88 PRIOR APPLICATION NUMBER: 60/090349
89 PRIOR FILING DATE: 1998-06-23
90 PRIOR APPLICATION NUMBER: 60/090355
91 PRIOR FILING DATE: 1998-06-23
92 PRIOR APPLICATION NUMBER: 60/090429
93 PRIOR FILING DATE: 1998-06-24
94 PRIOR APPLICATION NUMBER: 60/090431
95 PRIOR FILING DATE: 1998-06-24
96 PRIOR APPLICATION NUMBER: 60/090435
97 PRIOR FILING DATE: 1998-06-24
98 PRIOR APPLICATION NUMBER: 60/090444
99 PRIOR FILING DATE: 1998-06-24
100 PRIOR APPLICATION NUMBER: 60/090445
101 PRIOR FILING DATE: 1998-06-24
102 PRIOR APPLICATION NUMBER: 60/090472
103 PRIOR FILING DATE: 1998-06-24
104 PRIOR APPLICATION NUMBER: 60/090535
105 PRIOR FILING DATE: 1998-06-24
106 PRIOR APPLICATION NUMBER: 60/090540
107 PRIOR FILING DATE: 1998-06-24
108 PRIOR APPLICATION NUMBER: 60/090542
109 PRIOR FILING DATE: 1998-06-24
110 PRIOR APPLICATION NUMBER: 60/090557
111 PRIOR FILING DATE: 1998-06-24
112 PRIOR APPLICATION NUMBER: 60/090676
113 PRIOR FILING DATE: 1998-06-25
114 PRIOR APPLICATION NUMBER: 60/090678
115 PRIOR FILING DATE: 1998-06-25
116 PRIOR APPLICATION NUMBER: 60/090690
117 PRIOR FILING DATE: 1998-06-25
118 PRIOR APPLICATION NUMBER: 60/090694
119 PRIOR FILING DATE: 1998-06-25
120 PRIOR APPLICATION NUMBER: 60/090695
121 PRIOR FILING DATE: 1998-06-25
122 PRIOR APPLICATION NUMBER: 60/090696
123 PRIOR FILING DATE: 1998-06-25
124 PRIOR APPLICATION NUMBER: 60/090862
125 PRIOR FILING DATE: 1998-06-26
126 PRIOR APPLICATION NUMBER: 60/090863
127 PRIOR FILING DATE: 1998-06-26
128 PRIOR APPLICATION NUMBER: 60/091360
129 PRIOR FILING DATE: 1998-07-01
130 PRIOR APPLICATION NUMBER: 60/091478
131 PRIOR FILING DATE: 1998-07-02
132 PRIOR APPLICATION NUMBER: 60/091544
133 PRIOR FILING DATE: 1998-07-01
134 PRIOR APPLICATION NUMBER: 60/091519
135 PRIOR FILING DATE: 1998-07-02
136 PRIOR APPLICATION NUMBER: 60/091626
137 PRIOR FILING DATE: 1998-07-02
138 PRIOR APPLICATION NUMBER: 60/091633
139 PRIOR FILING DATE: 1998-07-02
140 PRIOR APPLICATION NUMBER: 60/091978
141 PRIOR FILING DATE: 1998-07-07
142 PRIOR APPLICATION NUMBER: 60/091982
143 PRIOR FILING DATE: 1998-07-07
144 PRIOR APPLICATION NUMBER: 60/092182
145 PRIOR FILING DATE: 1998-07-09

Query Match	99.3%;	Score 4719.5;	DB 9;	Length 867;
Best Local Similarity	99.5%;	Pred. No. 0;		
Matches 866;	Conservative 0;	Mismatches 1;	Indels 3;	Gaps 1;
Qy	1	MGPPSLVLCCLISATVFSLLGSSAFSLSHRLKGRFORDRRNIRPNIIILVLTDDQVGLS	60	
Db	1	MGPPSLVLCCLISATVFSLLGSSAFSLSHRLKGRFORDRRNIRPNIIILVLTDDQVGLS	60	
Qy	61	MQVMNKRRIIMEQGGTHFINAFVITPMCCPSRSSILTKGYVHNHTYTNNENCSPPSQWA	120	
Db	61	MQVMNKRRIIMEQGAHFINAFVITPMCCPSRSSILTKGYVHNHTYTNNENCSPPSQWA	120	
Qy	121	QHESTFVYLNSTGYRTAFGKYLNEVNGSVVPGWKWGLLKNRFFNYTLCRGVK	180	
Db	121	QHESTFVYLNSTGYRTAFGKYLNEVNGSVVPGWKWGLLKNRFFNYTLCRGVK	180	
Qy	181	EXHGSYSDKYLTDLITNDVSVSFFRTSKMYPHRPVLVMI SHAAPHGPDSPAQYSLRFP	240	
Db	181	EXHGSYSDKYLTDLITNDVSVSFFRTSKMYPHRPVLVMI SHAAPHGPDSPAQYSLRFP	240	
Qy	241	NASQHITPSYNAPNPDGHWIMRYTGPMPKPIHMEFTNMLQKRLOTILMSVDDSMETIYNM	300	
Db	241	NASQHITPSYNAPNPDGHWIMRYTGPMPKPIHMEFTNMLQKRLOTILMSVDDSMETIYNM	300	
Qy	301	LVETGELONTYIVYTADHGHIYHGFGLVKGKSMPEYFDIRVPFYVGRPNVEAGCLNPHIV	360	
Db	301	LVETGELONTYIVYTADHGHIYHGFGLVKGKSMPEYFDIRVPFYVGRPNVEAGCLNPHIV	360	
Qy	361	LNIDLAFTILDLAGLDIPADMDGKSLKLLDTERPVNRFHLKKMRVWRDSFLVERGKLL	420	
Db	361	LNIDLAFTILDLAGLDIPADMDGKSLKLLDTERPVNRFHLKKMRVWRDSFLVERGKLL	420	
Qy	421	HKRDNDKVDQAEENFLPKYQVKDLCQBAEYQTAQCEQLGQKWCQVEDATGKLHKCKGP	480	
Db	421	HKRDNDKVDQAEENFLPKYQVKDLCQBAEYQTAQCEQLGQKWCQVEDATGKLHKCKGP	480	
Qy	481	MRLGGSRALSNLVPKYGGGSEACTCDSDGYKLSLAGRRKKLFKKYKASTVRSRSIRSV	540	
Db	481	MELGGSRALSNLVPKYGGGSEACTCDSDGYKLSLAGRRKKLFKKYKASTVRSRSIRSV	540	
Qy	541	AIEVDGRVTHVGLGDAAPRNITKRWPCAPEDQDDKGGFSGTGGLPDYSAANPIKVT	600	
Db	541	AIEVDGRVTHVGLGDAAPRNITKRWPCAPEDQDDKGGFSGTGGLPDYSAANPIKVT	600	
Qy	601	HRCYILENTDVOCDLDLYKSLQAKDHLIDHEIETLQNKIKNLRVGRHLKKRPREEC	660	
Db	601	HRCYILENTDVOCDLDLYKSLQAKDHLIDHEIETLQNKIKNLRVGRHLKKRPREEC	660	
Qy	661	DCHKISYHTQHKGRUKHRGSSLHPRFKGLQEKVDKVLRLREQKRKKLKLRLQNNDTC	720	
Db	661	DCHKISYHTQHKGRUKHRGSSLHPRFKGLQEKVDKVLRLREQKRKKLKLRLQNNDTC	720	
Qy	721	SNPGLTCFTHDNOHWOTAPFWTLGPFCACTSANNTYMCRTINETHFLFCEPATGFL	780	
Db	721	SNPGLTCFTHDNOHWOTAPFWTLGPFCACTSANNTYMCRTINETHFLFCEPATGFL	780	
Qy	781	YFDLNTDTPQLMNAVNTLDRVLNLQVLMELRSCGYKQCNPRTRNMDLGLKDGGSYE	840	
Db	781	YFDLNTDTPQLMNAVNTLDRVLNLQVLMELRSCGYKQCNPRTRNMDL---DGGSYE	837	
Qy	841	QYRQFORRWKPEMRPSSKSLGQWEGWEG	870	
Db	838	QYRQFORRWKPEMRPSSKSLGQWEGWEG	867	

RESULT 12

US-09-989-731-84
; Sequence 84, Application US/09989731
; Patent No. US20020103125A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC70
CURRENT APPLICATION NUMBER: US/09/989,731
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202

;
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24

;
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 99.3%; Score 4719.5; DB 10; Length 867;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Qy/ 1 MGPPLVLCILSNVFSILGSSSAFLSHRLKGRFORDRNIRPNILVLTDDQDVELGS 60
Db 1 MGPPLVLCILSNVFSILGSSSAFLSHRLKGRFORDRNIRPNILVLTDDQDVELGS 60
Qy 61 MQVANKTRIMEQGGTHFINAFVTTMCCPSRSSILTKYVHNHTYTNNECSPSWSQA 120
Db 61 MQVANKTRIMEQGGTHFINAFVTTMCCPSRSSILTKYVHNHTYTNNECSPSWSQA 120
Qy 121 QHSRTFAVYLNSTGYRTAFFGKYLNEVNGSVPPGKWEVGLLKNRFPYNTLCRNGVK 180
Db 121 QHSRTFAVYLNSTGYRTAFFGKYLNEVNGSVPPGKWEVGLLKNRFPYNTLCRNGVK 180
Qy 181 EKHGSDYSKDYLTDLITNDVSFFRTSKKYPHPVLAVISHAAPHGSDSAPQYSRLFP 240
Db 181 EKHGSDYSKDYLTDLITNDVSFFRTSKKYPHPVLAVISHAAPHGSDSAPQYSRLFP 240
Qy 241 NASCHITPSYNYAPNPDKHWIMRYTGMKPIHMEFTNMLQKRLQTLASVDDSMETIYNM 300
Db 241 NASCHITPSYNYAPNPDKHWIMRYTGMKPIHMEFTNMLQKRLQTLASVDDSMETIYNM 300

```

QY 301 LVETGELONTYIVYTADHGYHIGQGLVKGKSMPEYFDIRPFFYVRGPNVEAGCLNPHIV 360
Db 301 LVETGELONTYIVYTADHGYHIGQGLVKGKSMPEYFDIRPFFYVRGPNVEAGCLNPHIV 360
QY 361 LNIIDLAPTILDIAGLDIPADMDGKSIILKLLDTERPVRNFHLKKGKVRWRDSDLVERGKLL 420
Db 361 LNIIDLAPTILDIAGLDIPADMDGKSIILKLLDTERPVRNFHLKKGKVRWRDSDLVERGKLL 420
QY 421 HKRDNKDVAEENFLPKYQRYKDLQRAEYQTAQELGQKQWQVEDATGKLKHLKCKGP 480
Db 421 HKRDNKDVAEENFLPKYQRYKDLQRAEYQTAQELGQKQWQVEDATGKLKHLKCKGP 480
QY 481 MELGSSRALSNLVPKYGGGSEACTCDSDYKLSLAGRRKLLPKKKYKASVYRSRSIRSV 540
Db 481 MELGSSRALSNLVPKYGGGSEACTCDSDYKLSLAGRRKLLPKKKYKASVYRSRSIRSV 540
QY 541 AIEVDGRVTHVGLGDAQAQPNRLTKRHWPACAPEDQDDKGGPFGSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVTHVGLGDAQAQPNRLTKRHWPACAPEDQDDKGGPFGSGTGGLPDYSAANPIKVT 600
QY 601 HRCYILENTVQCDLLDYKSLQAWKHKLHIDHEITETLQNKILNREVRGHLKXKRPBEC 660
Db 601 HRCYILENTVQCDLLDYKSLQAWKHKLHIDHEITETLQNKILNREVRGHLKXKRPBEC 660
QY 661 DCHKISYTHQHGRLKGRSSLHPFRKGLQEKDKVWLLREQKXKLLKRLQNNDDTC 720
Db 661 DCHKISYTHQHGRLKGRSSLHPFRKGLQEKDKVWLLREQKXKLLKRLQNNDDTC 720
QY 721 SMPGLTCFTHDNOHQWTAFTWTLGPFCACTSANNTNYWCMRTINETHNPLFCFATGPLE 780
Db 721 SMPGLTCFTHDNOHQWTAFTWTLGPFCACTSANNTNYWCMRTINETHNPLFCFATGPLE 780
QY 781 YFDLNTDTPQLMNAVNTLDRDVLNQLHVOLMELRCKGVKQCNPRTRNMDLGLKDGGSYE 840
Db 781 YFDLNTDTPQLMNAVNTLDRDVLNQLHVOLMELRCKGVKQCNPRTRNMDLGLKDGGSYE 840
QY 841 QYRQFQRRKWPENKRPSSKSLQQLMEGWEG 870
Db 838 QYRQFQRRKWPENKRPSSKSLQQLMEGWEG 867

```

RESULT 13

```

US-09-989-732-84
; Sequence 84, Application US/09989732
; Patent No. US20020123463A1
; GENERAL INFORMATION:
; APPLICANT: Aehkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; ACIDS Encoding the Same
; FILE REFERENCE: P2730P1C57

```

```

; CURRENT APPLICATION NUMBER: US/09/989,732
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876

```

; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25

; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 99.3%; Score 4719.5; DB 10; Length 867;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Qy 1 MGPPSLVLCILLSATVPSLLGSSAFSLSHRLKGRFORDRNRPNILVLTDQDVGLGS 60
Db 1 MGPPSLVLCILLSATVPSLLGSSAFSLSHRLKGRFORDRNRPNILVLTDQDVGLGS 60

Qy 61 MQVNMKTRIMEGGGTHFINAVFTTMCPCSSSILTGKYVHNHTYNNENCSSPSWQA 120
Db 61 MQVNMKTRIMEGGGAHPINAFVTTMCPCSSSILTGKYVHNHTYNNENCSSPSWQA 120

Qy 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKEMVGLLKNRFPYNTTLCRNGVK 180
Db 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKEMVGLLKNRFPYNTTLCRNGVK 180

Qy 181 EKHGSDYSKOYLTDLITNDOSVSFFRTSKQWPHRPVLMVISHAAPHGPDSDAPQYSRLFP 240
Db 181 EKHGSDYSKOYLTDLITNDOSVSFFRTSKQWPHRPVLMVISHAAPHGPDSDAPQYSRLFP 240

Qy 241 NASQHITPSYNTAPNDKHWIMRYTGPMPKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNN 300
Db 241 NASQHITPSYNTAPNDKHWIMRYTGPMPKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNN 300

Qy 301 LVETGELDNTYIVYTADHGYHIGQGLVKGKSNPYEFDIRVPFYVRGPNVEAGCLNPHIV 360
Db 301 LVETGELDNTYIVYTADHGYHIGQGLVKGKSNPYEFDIRVPFYVRGPNVEAGCLNPHIV 360

Qy 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRWVRDSEFLVERGKLL 420
Db 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKKRWVRDSEFLVERGKLL 420

Qy 421 HKRDNDKVDAAQENFLPKYQVRKDLQORAEYQACEOLGQKWCQVEDATGKLLHKCKGP 480
Db 421 HKRDNDKVDAAQENFLPKYQVRKDLQORAEYQACEOLGQKWCQVEDATGKLLHKCKGP 480

Qy 481 MRLGSSRALSNLVPKYYGQSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540
Db 481 MRLGSSRALSNLVPKYYGQSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540

Qy 541 AIEVDGRVTHVGLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVTHVGLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGGLPDYSAANPIKVT 600

Qy 601 HRCYILENDTVQCDLLYKSLQAKDHLHIHDETTLQNKIKNLRVGRHGLKKRPREC 660
Db 601 HRCYILENDTVQCDLLYKSLQAKDHLHIHDETTLQNKIKNLRVGRHGLKKRPREC 660

; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089908
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/089947
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089948
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/089952
 ; PRIOR FILING DATE: 1998-06-19
 ; PRIOR APPLICATION NUMBER: 60/090246
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090252
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090254
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090349
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090355
 ; PRIOR FILING DATE: 1998-06-23
 ; PRIOR APPLICATION NUMBER: 60/090429
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090431
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090435
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090444
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090472
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090535
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090540
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090542
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090676
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090678
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090690
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090862
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match 99.3%; Score 4719.5; DB 10; Length 867;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

 Qy 1 MGPPSLVLCLLSATVFSLLGGSSAFUSHHRLKGRFORDRNRPNILVLTDODVGLS 60
 Db 1 MGPPSLVLCLLSATVFSLLGGSSAFUSHHRLKGRFORDRNRPNILVLTDODVGLS 60

 Qy 61 MQVMNKTTRIMEGGTHFINAFVTTMCCPSRSSILTKYVHNHNTYTNNENCSSPSWQA 120
 Db 61 MQVMNKTTRIMEGGTHFINAFVTTMCCPSRSSILTKYVHNHNTYTNNENCSSPSWQA 120

 Qy 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKKEWVGLLKNRSFYNTYLCNGVK 180
 Db 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYNGSVYPPGKKEWVGLLKNRSFYNTYLCNGVK 180

 Qy 181 EKHSYDYSKDYLTDLTNDVSFFRTSKMYPHRPVLMVISHAAPHGSDSAPQYSLFP 240
 Db 181 EKHSYDYSKDYLTDLTNDVSFFRTSKMYPHRPVLMVISHAAPHGSDSAPQYSLFP 240

 Qy 241 NASQHITPSYNYAPNPKHWMRYTGPMKPIHMEFTNMLQKGLQTLMSVDDSMETIYNN 300
 Db 241 NASQHITPSYNYAPNPKHWMRYTGPMKPIHMEFTNMLQKGLQTLMSVDDSMETIYNN 300

 Qy 301 LVETGELDNITYIVYTADHGYHICQFGLVKGKSMPEYFDIRVPPYVGRPNVAGCLNPHIV 360
 Db 301 LVETGELDNITYIVYTADHGYHICQFGLVKGKSMPEYFDIRVPPYVGRPNVAGCLNPHIV 360

 Qy 361 LNIDLAPTILDIAGLDIPADMCKSILKLLDTERPVNRFHLKKKMRVWRDSDFLVERGKLL 420
 Db 361 LNIDLAPTILDIAGLDIPADMCKSILKLLDTERPVNRFHLKKKMRVWRDSDFLVERGKLL 420

 Qy 421 HKRDNDKVDQAEBNPLPKYQRYKDLCORAEYQACBQLGOKWQCVEDATGKLLHKCKGP 480
 Db 421 HKRDNDKVDQAEBNPLPKYQRYKDLCORAEYQACBQLGOKWQCVEDATGKLLHKCKGP 480

 Qy 481 MRIGGSRALSNLVPKYVGGSEACTCDSDYKLSLAGRKKLPKKYKASYVRSRSIRSV 540
 Db 481 MRIGGSRALSNLVPKYVGGSEACTCDSDYKLSLAGRKKLPKKYKASYVRSRSIRSV 540

 Qy 541 AIEVDGRVYHVGDAQAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
 Db 541 AIEVDGRVYHVGDAQAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600

 Qy 601 HRCYILENDTVQCDLDLYKSLQAWKDLHIHIEITLQNKIKNLRVGRHLKKGRPEEC 660
 Db 601 HRCYILENDTVQCDLDLYKSLQAWKDLHIHIEITLQNKIKNLRVGRHLKKGRPEEC 660

 Qy 661 DCHKISYHTOKHGRGLKRGSSLHPFRKGLQEKDKVLLREOKKKKLRKLLKRLQNNDT 720
 Db 661 DCHKISYHTOKHGRGLKRGSSLHPFRKGLQEKDKVLLREOKKKKLRKLLKRLQNNDT 720

 Qy 721 SMFGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNLFCEFATGFLE 780
 Db 721 SMFGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNLFCEFATGFLE 780

 Qy 781 YFDLNTDTPYOLMNAVNTLDRDVLNQLHVOLMELSCKYKQCNPRTRNMDLGLKDGGSYE 840
 Db 781 YFDLNTDTPYOLMNAVNTLDRDVLNQLHVOLMELSCKYKQCNPRTRNMDLGLKDGGSYE 840

 Qy 841 QYRQFORRWKPEMKRPSKSLQGLWEGWEG 870
 Db 838 QYRQFORRWKPEMKRPSKSLQGLWEGWEG 867

RESULT 15
 US-09-990-442-84
 ; Sequence 84, Application US/09990442
 ; Patent No. US20020132252A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Fertara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730PIC8
CURRENT APPLICATION NUMBER: US/09/990,442
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435

Search completed: October 15, 2003, 13:19:02
Job time : 83 secs

Query Match 99.3%; Score 4719.5; DB 10; Length 867;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2003, 13:04:35 ; Search time 26 Seconds
(without alignment)
3217.952 Million cell updates/sec

Title: US-09-970-287-2

Perfect score: 4751

Sequence: 1 MGPPSLVLCLSATVFSLLG.....PEMKRPSSKSLQLWEGWG 870

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1589	33.4	709	2	T16584
2	858.5	18.1	552	2	KJHUGU
3	453.5	9.5	639	2	T46577
4	452.5	9.5	649	2	S43229
5	341.5	7.2	646	1	KJUKM
6	302.5	6.4	497	2	G55169
7	301.5	6.3	497	2	C91206
8	301.5	6.3	497	2	E86052
9	290	6.1	535	2	AF0103
10	251	5.3	571	2	G90891
11	251	5.3	571	2	E49003
12	250.5	5.3	557	2	AF0370
13	250	5.3	571	2	B85726
14	246.5	5.2	517	2	AB0102
15	238.5	5.0	583	1	KJHUAC
16	238	5.0	551	2	S07089
17	238	5.0	567	2	A37362
18	234.5	4.9	550	1	KJHUID
19	233	4.9	497	2	AF0506
20	232	4.9	551	2	S01793
21	228	4.8	538	2	F83354
22	226	4.8	563	2	A47153
23	224	4.7	465	2	F70837
24	222	4.7	503	2	E83642
25	221	4.7	514	2	T44602
26	219.5	4.6	479	2	C75099
27	219	4.6	787	2	B70643
28	218.5	4.6	522	1	KJHUG6
29	211.5	4.5	533	1	KJHUAB

30	210.5	4.4	506	2	A54190	cerebroside-sulfat
31	209	4.4	503	2	AD3128	choline sulfatase
32	209	4.4	503	2	E98159	choline sulfatase
33	207	4.4	603	2	F70651	probable atsf prot
34	205.5	4.3	521	2	H87394	sulfatase family p
35	203	4.3	535	1	A44475	N-acetylgalactosam
36	203	4.3	577	2	T45548	arylsulfatase (EC
37	201.5	4.2	464	2	B35159	arylsulfatase (EC
38	200	4.2	593	2	I37186	arylsulfatase D (E
39	194.5	4.1	507	1	KJHUAA	cerebroside-sulfat
40	188	4.0	495	2	AD3523	choline-sulfatase
41	183.5	3.9	589	2	I37187	arylsulfatase E (E
42	181.5	3.8	555	2	D95377	probable cerebroso
43	178.5	3.8	451	2	G84166	hypothetical prote
44	176	3.7	551	2	S36691	arylsulfatase (EC
45	176	3.7	551	2	C91220	arylsulfatase [imp

ALIGNMENTS

RESULT 1

T16584

hypothetical protein K09C4.8 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C;Accession: T16584

R;Du, 2.

submitted to the EMBL Data Library, December 1995

A;Description: The sequence of C. elegans cosmid K09C4.

A;Reference number: Z18542

A;Accession: T16584

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-709 <DUZ>

A;Cross-references: EMBL:U43375; NID:g1125841; PID:g1125842; PIDN:AAA83618.1; CESP:K09C

C;Genetics:

A;Gene: CESP:K09C4.8

A;Introns: 33/2; 72/2; 110/2; 180/3; 237/2; 287/3; 346/2; 406/2; 448/3; 545/2; 564/2; 1

Query Match 33.4%; Score 1589; DB 2; Length 709;

Best Local Similarity 40.2%; Pred. NO. 9.3e-107;

Matches 317; Conservative 118; Mismatches 196; Indels 158; Gaps 17;

QY	45	NILVLTDQDDVGLSGVMVNTTRIMEQGGTHFNAFVTTMCCPSSSILITGKYVNH	104
DB	36	NVLILTDQDDVGLSGMDPMTSQIMKRGTEFTSGYVTTTFCPCRSIITGLYVNH	95
QY	105	NTYTNNCCSPSQAHESRTFAVYLNSTGYRTAPFGKYLNEYSGYVPPQKMWGLL	164
DB	96	HVHTNNQNTGVEWRKVHEKKSIGVYVLEAGYRTAYLGKYLNEYDGSYIPQGDWHAI	155
QY	165	KNSRFNYTLCRNGVKEKGSYSDYLDLTNTDSVSFFRTSKMYPHRPVLMVISHAA	224
DB	156	KNSKFNYTMNGEREKFGSEYKDYFDLTNRSKLFIDKHIRAWQPPALLISYPA	215
QY	225	PHGPEDSAPOYSRLFPNASOHIPTPSYVAPNDPKWIMRYTGPMPKPIHMEFTNMLQRKL	284
DB	216	PHGPEDPAQPAHMEFENEISHTGWNFAPNDPKWLLQRTGQNDVHISFTDLHRRRL	275
QY	285	QTLMSVDDSMETIYNMLVETGELDNTIYVYTDHGYHIGQFGLVKGSKMPYEFDIRVPPY	344
DB	276	QTLQSVDEGIERLFLNLLRELNLWNTYAIYTSDHGYHLGQFGLLKGMKMPYEFDIRVPPF	335
QY	345	VKGPVVEAGCLNPHIVNLIDAPTILDTAGLDIPADMGSKSLKLLDTERPVRPHLKKK	404
DB	336	MKGPGIPRNVTEINVTNVDIAPTMLHIAGVPKARNNGRSLLLEVALKKK-----KKK	389
QY	405	----MEVWEDSFLVGRGLKH-KRDNKDVADEENFLPKYQVVKDLCORAEVQTACEQLG	459
DB	390	HMTALKPMRDTTILIBRGKMPKJKTIRDIYIKQKKFP-NKENRLSKCKRKRQDQVH-G	447
QY	460	QKWCQVEDATGKLKHLKCKGPMRLGSRALSNLVPKYVQGSSEACTCDSDGYKLGLAGRR	519

Db 448 QAKCYTVEDRWRYKCR-----DNWSDQSC-----R 476
QY 520 KKLKKKKYKASVRSRSIRSAIEVDGRVYHVLGDAAQPRNLTKRWHPGAPEDQDDXG 579
Db 477 KK-----REI-----SNYDDDDI 489
QY 580 GDFSCTGGLPDYSAANPIKVTTHRCYILENDTVQCDLDLYKSLQAMKDKHLHIDHIEITLQ 639
Db 490 DEF-----LTYADRENFSGEHWTQGEFEDSGEV-----G 519
QY 640 NKIKNLRVGRHLKKRPECCDCHKISYHTQKGLKRGSSLSLHPRFKGLQEKQKVMLLR 699
Db 520 EELDGHRKRGILSK-----CSCSR-----NVSHPIK-----LLE 549
QY 700 EQRKRLKRLKRLKRLQND-----TCSMGLCTCTHDNQHWTAPFW-----TLGPPFACTSAN 753
Db 550 QKMSKHLKYLKYPKQNGSLKPKDCSLPOMNCFTHTASHMKTPPLWPELGEFCFCQCN 609
QY 754 NNTYCMRTINETHNLFCEPATGFEVFDLNTDPVQLMNAVNTLDRVLNQLHVLQML 813
Db 610 NNTYCLRTKNTNTHNLFCEFTVFISFYDFNTDPQLINAVISLDIGVLSQLRNL 669
QY 814 RSCGYKQC 822
Db 670 RKCKN-RQC 677
RESULT 2
KUHUGU
N-acetylglucosamine-6-sulfatase (EC 3.1.6.14) precursor - human
N/Alternate names: Chondroitin sulfatase, N-acetyl-D-glucosamine-6-sulfate 6-sulfohydrolase
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 27-Oct-1995 #text_change 24-Sep-1999
C:Accession: S27164; A31672
R:Robertson, D.A.; Freeman, C.; Morris, C.P.; Hopwood, J.J.
Biochem. J. 288, 539-544, 1992
A:Title: A cDNA clone for human glucosamine-6-sulphatase reveals differences between aryl
A:Reference number: S27164; MUID:93098807; PMID:1463457
A:Accession: S27164
A:Molecule type: mRNA
A:Residues: 1-552 <ROB>
A:Cross-references: GB:Z12173; EMBL:M23657; NID:G31866; PIDN:CAA78164.1; PID:G31867
A:Note: Parts of this sequence, including the amino end of the mature protein, were deter
A:Note: A form is described with a proteolytic cleavage releasing residue 222 as the ami
R:Robertson, D.A.; Freeman, C.; Nelson, P.V.; Morris, C.P.; Hopwood, J.J.
Biochem. Biophys. Res. Commun. 157, 218-224, 1988
A:Title: Human glucosamine-6-sulfatase cDNA reveals homology with steroid sulfatase.
A:Reference number: A31672; MUID:89061714; PMID:3196333
A:Accession: A31672
A:Molecule type: mRNA
A:Residues: 178-552 <RO2>
A:Cross-references: GB:Z12173; EMBL:M23657; NID:G31866
C:Genetics:
A:Gene: GDB:GNS
A:Cross-references: GDB:120006; OMIM:252940
A:Map position: 12q14-12q14
A:Note: Defects in this gene can cause mucopolysaccharidosis type III D, Sanfilippo D di
C:Function:
A:Description: hydrolyzes N-acetyl-D-glucosamine 6-sulfate units in heparan sulfate and
C:Superfamily: animal sulfatase
C:Keywords: glycoprotein; lysosomal storage disease; lysosome; Sanfilippo disease; sulfu
F:1-43/Domain: signal sequence #status predicted <SIG>
F:44-552/Product: N-acetylglucosamine-6-sulfatase #status predicted <MAT>
F:91/Modified site: 3-oxoalanine (Cys) #status predicted
F:111,117,183,198,210,279,317,362,387,405,449,480/Binding site: carbohydrate (Asn) (cova
F:422/Binding site: carbohydrate (Asn) (covalent) #status experimental
Query Match 18.1%; Score 858.5; DB 1; Length 552;
Best Local Similarity 27.2%; Pred. No. 4.4e-54;
Matches 256; Conservative 76; Mismatches 174; Indels 355; Gaps 18;
QY 4 PSILVCLLSAT--VPSLLGSSAFLSHRLKGRFQDRNRTRPNILVLTDDQDVELGSM 61

Db 24 PALLLVGLGCLGVFGVAGTR-----RPNVLLITDDQDEVLGGM 64
QY 62 QVNMKTRRIMEQGTTHFINAFVTTMCCPSRSRSLTGKTVHNTYTN--NENCSFSPSQ 119
Db 65 TPLKTKKALIGEMGTFSSAYVPSALCCPSRASILTGYPHNHVNVNTLEGNCSSKQ 124
QY 120 AQHESHTPAVYLS--TCYRTAFPGKYLNEYG-----SYVPGWKWVGLLKNSRFNY 172
Db 125 KIQEPNTPFAILASMGYQTFAGKYLYNEYGAPDAGGLEHVLGWSYVALEKNSKYNY 184
QY 173 TLRNGVKEKHGSDYSKOYLTDLITNDSVSFFRTSKMYPHRPVLMAVISHAAPHGEDSA 232
Db 185 TSLNGKARKGNGENYSDYLTDLVLANVSLDFDYKSNF--EPFFMMIATPAPHSWTA 241
QY 233 PQVSRUPPNASQHTPSYVAPNDKHWMT--GPMKPIHMEFTNMLQKRLQTLMSVD 291
Db 242 PQYQKAFQVVPAPRNKNFN--IHGTNKHMLIROAKTPMTNSSIQFLDNAFRKRWQTLISVD 300
QY 292 DSMETIYNMLVETGELDNTRYIVYTDHGYHIGOPGLVKCKSMPEYEDIRVPFYVRGPVNE 351
Db 301 DLVEKLVKRLVETGELDNTRYIVYTDHGYHIGOPGLVKCKSMPEYEDIRVPFYVRGPVNE 360
QY 352 AGCINPHIVNLIDLAFTILDIAGLDI--PADMDGKSLKLLDTERPVNRPFLKKKQVWRD 410
Db 361 PNQTSKMLVANIDLGPTILDIAGYDLNKTQMDGMSLLPIL---RGASNL-----TWRS 410
QY 411 SFLVERKLLHKKRDNKVDQAQENFLPKYQVKNLCORAEYQVACQOLGKQKQVEDATG 470
Db 411 DVLVE----- 415
QY 471 KKLHKCKGPMRLGSSRALSNLVPKYVGSGSEACTCDSDGYKLSLAGRRKKLFXKXKAS 530
Db 416 -----YQSG----- 420
QY 531 YRSRSIRSAIEVDGRVYHVLGDAAQPRNLTKRWHPGAPEDQDDKOGDGFSGTGLPD 590
Db 421 -----RNV----- 424
QY 591 YSAANPIKVTTHRCYILENDTVQCDLDLYKSLQAMKDKHLHIDHIEITLQNKIKNLRVGR 650
Db 425 ----- 424
QY 651 HLKKRPEECDCCHKISYHTQKGLKRGSSLSLHPRFKGLQEKQKVMLLREQKRLKRL 710
Db 425 -----DPTCPSL----- 432
QY 711 LKRLQNNDTCSMPGLT-CPTHNDQHWQVAPFWTLGPFCACTSANNTYWMRTINETHNF 769
Db 433 -----PGVSQCF-----PDCVCEADAYNNTYACVRTMSALMNL 464
QY 770 LFCEP--ATGFLYEPDLNTPVQLMNAVNTLDRVLNQLHVLQMLRSCG 818
Db 465 QYCFDDQEVFVEVYNTADDPDQITNIANTIDPELLGQWYRLMMLQSCG 515
RESULT 3
T46577
aryl sulfatase (EC 3.1.6.1) [validated] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 04-Mar-2000
C:Accession: T46577
R:Baker, D.L.; Paletta, J.V.
submitted to the EMBL Data Library, February 1997
A:Description: Molecular characterization of the aryl sulfatase gene of Neurospora cras
A:Reference number: Z23090
A:Accession: T46577
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-639 <BAK>
A:Cross-references: EMBL:U89492; PIDN:AAC02716.1
A:Experimental source: strain wild type 74-OR23-1A
C:Genetics:

A:Gene: ars-1
A:Map position: 7
A:Introns: 115/1; 518/3
C:Function:
A:Description: EC 3.1.6.1 [validated, MUID:89384583]
A:Note: genes are expressed under conditions of sulfur limitation and are under coordination
C:Superfamily: plant sulfatase
C:Keywords: sulfuric ester hydrolase

Query Match	9.5%	Score 453.5	DB 2	Length 639
Best Local Similarity	27.7%	Prod No. 1.e-24		
Matches 139	Conservative	73	Mismatches 211	Indels 77
				Gaps 14
QY	13	ATVPSLLGGSSAFLSHRLKGRFQRDRNRIRPNILVTDQDQVGLSGMQVMKTRIME	72	
DB	15	STAPADVSGQSGFGEHKPKGTTEGKK-SPNIVFILTDDQLHLQDLYLPLKKYLA	73	
QY	73	QCGTHFINAFVTTPMCCPSSSILTKYVHNHN-TYTNNENCSPPSMAQHEKRTF-AVY	130	
DB	74	DEGTTYKRHYCTTACCPCARVSLWTQKAHNTNVDVSPYGGYKPKFISQGFNEAYLPVW	133	
QY	131	LNSGTGRTAPFGKYLN-----EYNGSVYPPGWKEWGLLK--NSRFNYTYLCRNGVKEKH	183	
DB	134	LQKAGYDTTYTGKLFNAHTVDNDSPIY-AGWNGSDFLLDPTYTYSYLNATFORN---RDP	189	
QY	184	GSDYSKDYLTLITDNDVSFRTSKQYPRPVLWLVISHAAPHG-----	227	
DB	190	PISYEGQYSDVLAEXAYGFLDEAAKNVNRPFELGIAPLAPHSNVEPGFPSSSSSSSS	249	
QY	228	-----PESSAPQYSLFPNASQHIPTSYNYANPD-KIWIWR	263	
DB	250	DSATLHRRPTNEHDDIEKVSFTPTPAARAHALFPDVIIVRTPHENPSSRAGSVMIAR	309	
QY	264	YTGMPKPIHNEFTMQLQRKLOTLMSVDDSMETIYNMLVETGELDNITYIYTTADHGTHIG	323	
DB	310	LP-HQSAENVAFNDFHFRQRLRALESVDLAVAGVVARLERHGLLENTYVFTYADNGYHIG	368	
QY	324	QFGLVKGSMPEYFDIRVPYFVPGNVEACCLNPHVINLIDLAPTILDIAGL-----	375	
DB	369	QHRLOPQKECGFEEDINVLIVRPGYKKEVAEVTSVSHVDLAPTIVLKLALAGLDKESKE	428	
QY	376	---DIPADMGKSI---LKLILDTERPVNRPHLKKKRWVSDSFLVERGKLLHKRDNKDVD	429	
DB	429	EEEBEYGLDGEAIPVTAEDLAQAKAKGKGEHVTVEYW--GPAVSGRVEFK-----D	480	
QY	430	AQENFLPKYORVKDLQOR	448	
DB	481	ATRYTWNTNTYKAVKIIGER	499	

C;Date: 28-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 18-Aug-2000
C;Accession: S43229; S68892; S41433
R;Hallmann, A.; Sumper, M.
E;Eur. J. Biochem. 221, 143-150, 1994
A;Title: An inducible arylsulphatase of Volvox carteri with properties suitable for a rep-
A;Reference number: S43229; MUID:94222055; PMID:8168504
A;Accession: S43229
A;Molecule type: mRNA
A;Residues: 1-649 <HAL>
A;Cross-references: EMBL:X77214; NID:9452338; PIDN:CAA54426.1; PID:9452339
A;Note: Part of this sequence, including the amino end of the mature protein was confir-
R;Selmer, T.; Hallmann, A.; Schmidt, B.; Sumper, M.; von Figura, K.
E;Eur. J. Biochem. 238, 341-345, 1996
A;Title: The evolutionary conservation of a novel protein modification, the conversion o-
A;Reference number: S68892; MUID:96283826; PMID:8681943
A;Accession: S68892
A;Molecule type: protein
A;Residues: 64-71, 'S', 73-76 <SEL>
C;Function:

C;Keywords: glycoprotein; sulfuric ester hydrolase

F;1-21/Domain: signal sequence #status predicted <SIG>

F;22-646/Product: arylsulfatase #status predicted <HAT>

F;42,443/Binding site: carbonylate (Asn) (covalent) #status predicted

F;73/Modified site: 3-oxoalanine (Cys) #status predicted

Query Match 7.2%; Score 341.5; DB 1; Length 646;

Best Local Similarity 24.6%; Pred. No. 1.4e-16;

Matches 124; Conservative 72; Mismatches 180; Indels 129; Gaps 18;

QY 43 RPNILVLDDDDQDELGSM--QVMKTRIMEQGGTHFINAFVTPMCCPSRSSILTCKY 100

DB 26 KNFVVIITDDQDAIONSTHYPMSLHKYIRYGVLSQYFVTPVPCPSRTNLCAAS- 84

QY 101 VHNHTYTNENC-----SPSQAQHSR-----TFAYVLSNCGVTRTAFF 141

DB 85 --SPTFTSPACCLPTVAGSGRAWASTPTCRGSRITKAITPTTWASSLMTPTSATTSR 142

QY 142 GKVLNYSYVPPGKWEVGLLKNSRFNYTLCRNGVKEKHGSDYSKDYLTDLITDSV 201

DB 143 CRGLGRYRCVTPYTFD-----YNTLQNGATP---NYPGEYSTDVIRKGV 188

QY 202 SPFRTSKQMPHRPVLWVISHAAPHGPDSPAPQYSLRFPNASCHITPSYNAFNP-DRHW 260

DB 189 AOKSA--VAAGKFFAQISPIAPH-----TSQISTNPATGVTRSYFFPPIAPPWH 239

QY 261 IM-----RTGPMKPIHMEFTNMLQRLQTLMSVDD 292

DB 240 QMFSDANLPGGSOEPLRGRRQARLDPRPAAGPAEQPHLPGGLPPAP--EVAGPVDE 297

QY 293 SMETIYNMLVETGELNITYIVTADGHYHIGQFGLVKGKSMPEYFDIRVPFVVRGNVEA 352

DB 298 LIEQVVKTLDEAGVLDNTYIIISADNGYVGHARGAGKATCYEEDLAVPLIRGFIKA 357

QY 353 G-----CLNPHIVNLDTILDIAG-----LDIPADMCKSILKLLDTER 394

DB 358 SKSDKPKQNSKVLGVHDFAPITILSLAGASHLLGDKGLDGTPLGLYANDDG-----TLPSDY 412

QY 395 PVNRFLHKKMRV---WRDSFLVERGKLLHKKRDNKDQDAQENFLPKYQKVDLCQRAE 450

DB 413 PRPEQRQOFQGEFGWGSDELL-----QNRSQPNNTW-KVRYTD----- 453

QY 451 YOTACEQLQGRQCVEDATGKLKLH 475

DB 454 -----ESSQGWKLIQAQCTNERELY 473

RESULT 6

G65169

probable sulfatase (EC 3.1.6.-) - Escherichia coli (strain K-12)

N;Alternate names: hypothetical 57.3K protein ilvO-1bpB intergenic region

C;Species: Escherichia coli

C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002

C;Accession: G65169

R;Blattner, F.J.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: G65169

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-497 <BLAT>

A;Cross-references: GB:AE000445; GB:U00096; NID:gl1790105; PIDN:AACT6761.1; PID:gl1790112;

A;Experimental source: strain K-12, substrain MGL655

C;Genetics:

A;Gene: yidJ

C;Keywords: sulfuric ester hydrolase

F;52/Modified site: 3-oxoalanine (Cys) #status predicted

Query Match 6.4%; Score 302.5; DB 2; Length 497;

Best Local Similarity 25.5%; Pred. No. 6.2e-14;

Matches 105; Conservative 72; Mismatches 154; Indels 81; Gaps 21;

QY 43 RPNILVLDDDDQDELGSM--QVMKTRIMEQGGTHFINAFVTPMCCPSRSSILTCKY 97

DB 3 RPNFLPMTDQATNNMVGCGYSGKPLN-TQIDSLAAEGIRFNSAVTCSVCTPARAGLFT 61

QY 98 GKVVHNHTYTN--ENCSSPSWQAQHSRTPAVYLNSTGYRTAFPGKYL---NEYNG 150

DB 62 GIYANOSGPMNTNNAVPGKNIS-----TMGRYFKDAGYHTCYTGKWHLDGHDYFG 110

QY 151 S-YVPPGK-----EWVGLLKNSRFNYTLCRNGVKEKHGSDYSKDYLTDL-----ITN 198

DB 111 TGCCPPEMDADYWFDCANLYSELTEKISLWENGLNSV--EDLQANHIDEITFTWAHRISN 168

QY 199 DSVSFRTSKQMPHRPVLWVISHAAPHGP-----EDSAPQYSRLFPNNAHQHITPSY 250

DB 169 RAVDFLQOPAR--AEEFFLMVSYDEPHHPFTCPVEYLEKYADFYELGKQAQDDL--- 222

QY 251 NYAPNPKH--WIMRYTGP-----KPIHMEFTNMLQRLQTLMSVDDSMETIYNML 301

DB 223 --ANKPEHRLWAQAMPSPVGGDGLYHHPLYFACNDF-----VDDQIGRVINAL 269

QY 43 RPNILVLDDDDQDELGSM--QVMKTRIMEQGGTHFINAFVTPMCCPSRSSILTCKY 97

DB 3 RPNFLPMTDQATNNMVGCGYSGKPLN-TQIDSLAAEGIRFNSAVTCSVCTPARAGLFT 61

QY 98 GKVVHNHTYTN--ENCSSPSWQAQHSRTPAVYLNSTGYRTAFPGKYL---NEYNG 150

DB 62 GIYANOSGPMNTNNAVPGKNIS-----TMGRYFKDAGYHTCYTGKWHLDGHDYFG 110

QY 151 S-YVPPGK-----EWVGLLKNSRFNYTLCRNGVKEKHGSDYSKDYLTDL-----ITN 198

DB 111 TGCCPPEMDADYWFDCANLYSELTEKISLWENGLNSV--EDLQANHIDEITFTWAHRISN 168

QY 199 DSVSFRTSKQMPHRPVLWVISHAAPHGP-----EDSAPQYSRLFPNNAHQHITPSY 250

DB 169 RAVDFLQOPAR--AEEFFLMVSYDEPHHPFTCPVEYLEKYADFYELGKQAQDDL--- 222

QY 251 NYAPNPKH--WIMRYTGP-----KPIHMEFTNMLQRLQTLMSVDDSMETIYNML 301

DB 223 --ANKPEHRLWAQAMPSPVGGDGLYHHPLYFACNDF-----VDDQIGRVINAL 269

QY 302 VERGELDNTYIVYVADGHYHIGQFGLV-KGKSMPEYFDIRVPFVVRGNVEAGCLNPHIV 360

DB 270 TPE-QRENTWIVYTSDHGEMGAHKLISKGAAM-YDDITRIPLIIRSPQGERRQVDT-P 326

QY 361 LNTDLAPTILDTAGLDIPADMCKSILKLLDTERPVNRFLHKKRWRDVSF 412

DB 327 SHIDLLPTMALADIEKPEILFGENILAVKEPRGVMEFN---RYEIHDSF 375

RESULT 7

C91206

probable sulfatase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05099

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001

C;Accession: C91206

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.;

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and ge

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: C91206

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-497 <HAY>

A;Cross-references: GB:BA000007; PIDN:BA838042.1; PID:gl3364094; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: Ece4619

Query Match 6.3%; Score 301.5; DB 2; Length 497;

Best Local Similarity 25.5%; Pred. No. 7.4e-14;

Matches 105; Conservative 73; Mismatches 153; Indels 81; Gaps 21;

QY 43 RPNILVLDDDDQDELGSM--QVMKTRIMEQGGTHFINAFVTPMCCPSRSSILTCKY 97

DB 3 RPNFLPMTDQATNNMVGCGYSGKPLN-TQIDSLAAEGIRFNSAVTCSVCTPARAGLFT 61

QY 98 GKVVHNHTYTN--ENCSSPSWQAQHSRTPAVYLNSTGYRTAFPGKYL---NEYNG 150

DB 62 GIYANOSGPMNTNNAVPGKNIS-----TMGRYFKDAGYHTCYTGKWHLDGHDYFG 110

QY 151 S-YVPPGK-----EWVGLLKNSRFNYTLCRNGVKEKHGSDYSKDYLTDL-----ITN 198

DB 111 TGCCPPEMDADYWFDCANLYSELTEKISLWENGLNSV--EDLQANHIDEITFTWAHRISN 168

QY 199 DSVSFRTSKQMPHRPVLWVISHAAPHGP-----EDSAPQYSRLFPNNAHQHITPSY 250

DB 169 RAVDFLQOPAR--AEEFFLMVSYDEPHHPFTCPVEYLEKYADFYELGKQAQDDL--- 222

QY 251 NYAPNPKH--WIMRYTGP-----KPIHMEFTNMLQRLQTLMSVDDSMETIYNML 301

DB 223 --ANKPEHRLWAQAMPSPVGGDGLYHHPLYFACNDF-----VDDQIGRVINAL 269

$\frac{1}{\sqrt{\pi}}$

90

Db 66 DD L G Y Q L P F D K T S F D P K S M E D R D V V D T Y K I G I D K A I E A A K S T P T L L S L M D E - G V R L T N

[illegible]

Search completed: October 15, 2003, 13:09:49
Job time : 30 secs

Search completed: October 15, 2003, 13:09:49
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2003, 13:01:00 ; Search time 15 Seconds

(without alignments)
2727.549 Million cell updates/sec

Title: US-09-970-287-2

Perfect score: 4751

Sequence: 1 MGPPSLVCLLSATVFSLLG.....PEMKRPSSKSLGOLWEGWG 870

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	858.5	18.1	552	1	GL6S_HUMAN
2	856	18.0	559	1	P15866 capra hircu
3	452.5	9.5	649	1	GL6S_CAPIH
4	405	8.5	647	1	ARS_VOLCA
5	302.5	6.4	497	1	ARS_CHLRE
6	256	5.4	512	1	YIDJ_ECOLI
7	251	5.3	560	1	BETC_RHIME
8	238.5	5.0	583	1	YDEN_ECOLI
9	238	5.0	567	1	ARS_STRPU
10	234.5	4.9	550	1	ARS_STRPU
11	232	4.9	551	1	IDS_HUMAN
12	226	4.8	563	1	IDS_HEMPU
13	218.5	4.6	522	1	IDS_MOUSE
14	210.5	4.4	506	1	GA6S_HUMAN
15	210.5	4.4	533	1	ARS_MOUSE
16	210.5	4.4	577	1	ARS_MOUSE
17	203	4.3	535	1	STS_RAT
18	201.5	4.2	464	1	ARS_FELCA
19	200	4.2	593	1	ARS_KLEAE
20	198.5	4.2	591	1	ARS_HUMAN
21	194.5	4.1	507	1	ARS_HUMAN
22	189.5	4.0	624	1	ARS_MOUSE
23	183.5	3.9	589	1	STS_MOUSE
24	176.5	3.7	502	1	ARS_HUMAN
25	176	3.7	551	1	SPHM_HUMAN
26	170	3.6	473	1	ARS_ECOLI
27	155.5	3.3	535	1	ARS_RAT
28	138.5	2.9	510	1	ARS_PSEAE
29	137.5	2.9	510	1	GPMI_LISIN
30	135.5	2.9	510	1	GPMI_LISNO
31	134	2.8	533	1	GPMI_ANTSP
32	127	2.7	285	1	GPMI_ANASP
33	127	2.7	532	1	ARS_MOUSE
					P74507 synechocyst

34	125.5	2.6	724	1	ATIL_VACCV	P24759 vaccinia vi
35	124	2.6	510	1	GPMI_BACHD	Q9K716 bacillus ha
36	124	2.6	512	1	GPMI_OCEIH	P53174 oceanobacil
37	123	2.6	638	1	YOGS_BACSU	P54496 bacillus su
38	123	2.6	2339	1	RPL1_PLAPA	P27625 plasmodium
39	122	2.6	510	1	GPMI_CLOAB	Q97153 clostridium
40	122	2.6	534	1	GPMI_PORPU	P51379 porphyra pu
41	118.5	2.5	875	1	DD10_HUMAN	O13206 homo sapien
42	118	2.5	510	1	GPMI_BACST	Q9X519 bacillus st
43	117	2.5	512	1	GPMI_CLOPE	Q8XKU2 clostridium
44	112.5	2.4	1435	1	EBAL_PLAFC	P19214 plasmodium
45	112	2.4	1616	1	VITI_CAEEL	P55155 caenorhabdi

ALIGNMENTS

RESULT 1
GL6S_HUMAN
ID GL6S_HUMAN STANDARD; PRT: 552 AA.
AC P15866;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE N-acetylglucosamine-6-sulfatase precursor (EC 3.1.6.14) (G6S)
DE (Glucosamine-6-sulfatase).
GN GNS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Endothelial cells;
RX MEDLINE=93098807; PubMed=1463457;
RA Robertson D.A., Freeman C., Morris C.P., Hopwood J.J.;
RT "A cDNA clone for human glucosamine-6-sulphatase reveals differences
RT between arylsulphatases and non-arylsulphatases.";
RL Biochem. J. 288:539-544 (1992).
RN [2]
RP SEQUENCE OF 178-552 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=89061714; PubMed=3196333;
RA Robertson D.A., Freeman C., Nelson P.V., Morris C.P., Hopwood J.J.;
RT "Human glucosamine-6-sulfatase cDNA reveals homology with steroid
RT sulfatase.";
RL Biochem. Biophys. Res. Commun. 157:218-224 (1988).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 6-sulfate group of the N-
CC acetyl-D-glucosamine 6-sulfate units of heparan sulfate and
CC keratan sulfate.
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- DISEASE: DEFECTS IN GNS RESULT IN THE ACCUMULATION OF PARTIALLY
CC DEGRADED HEPARAN SULFATE IN LYSOSOMES CAUSING ORGANELLE CELL AND
CC TISSUE DISTORTION, ULTIMATELY LEADING TO THE LYSOSOMAL STORAGE
CC DISORDER, MUCOPOLYSACCHARIDOSIS TYPE IIID (MPS-IIID) (ALSO KNOWN
CC AS SANFILIPPO D SYNDROME).
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL: Z12173; CAA78164.1; -;
CC PIR: S27164; KJHUGU.
CC Genew: HGNC:4422; GNS.
CC MIM: 252940; -;
CC GO: GO:0008449; F:N-acetylglucosamine-6-sulfatase activity; TAS.
CC GO: GO:0006027; F:glycosaminoglycan catabolism; TAS.
CC InterPro: IPR000917; Sulfatase.
CC Pfam: PF00884; Sulfatase; 1.


```
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 559 AA; 62711 MW; 786CCDC48334A458 CRC64;

Query Match
Best Local Similarity 27.5%; Score 856; DB 1; Length 559;
Matches 231; Conservative 77; Mismatches 177; Indels 354; Gaps 20;

QY 3 PPSLVLCILSATVSLGGSSAFSLSHRLKGRFDRNRIRPNILVLTDQDVELGSMQ 62
Db 29 PPPLLLLL-----LGGCLG-----VSGAAGSR---RPNVVLVLAQQDEVLGGMT 72
QY 63 VMKTRIMEQGGTHFINAFVTPMCCPSRSSILTKYVHNHTYTN--NENCSSPSWQA 120
Db 73 PLKTKALIGEMGMFTFSAYVPSALCCPSRASILTKYVHNHTYTN--NENCSSPSWQA 132
QY 121 QHESRTAVYVLS--TGRTAFPGKYINEYNG-----SYPPGKWEVGLLKNRFRNYT 173
Db 133 IQEPNTFPAILRSMCGYQTFAGKYLNEYGAPDAGGLHVPGLGWSYWYALEKNSKYNT 192
QY 174 LCRNGVKEKGDYSKDVLTDLTINDSVSF--PRTSKMYPHRPVLMVISHAAPHGPEDS 231
Db 193 LSIKGARKGENTSVDTLTDVLNVLDFDYKNSN-----PFFMISTPAPHSWTA 247
QY 232 APQYSLFPNASQHITPSYNAFNPDKHWMRYT--GPMKPIHMEFTNMLQKRLTLMV 290
Db 248 APQYONAFQNVFAPRNKFN--IHGTNKHLLIRQAQKTPMNSSIQFLDNAPRERQWTLSS 306
QY 291 DDSNETYVNLVGEGLDNYIVTADHGYHIGQFGLVKGSMPIEDIRVYFVQGNV 350
Db 307 DDLVEKLKRLKPELNGELNNTYIFTSNGVHTGQFSLPIDKRLQYEDFDKVLVLRGPGI 366
QY 351 EAGCLNPHVNLIDAPLTDIAGLDI--PADMDCKSLKLLDTERPVNRPHLKKQKQVVR 409
Db 367 KPNQTSKMLVANILGFTLIDIAGYLNTKQMGNSFLPIL--RGASNL-----TWR 416
QY 410 DSFLVVERKLLHKRDNPKVDAQEENFLPKYQVXKDCQRAEYOTACBQLGQKQVCVEDAI 469
Db 417 SDVLVE-----RNV----- 422
QY 470 GKLLKHKCKGPMRLGGSEALSNLVPKYVGQSEACTCDSDGYKLSLAGRKKLFKKKYKA 529
Db 423 -----YQEG----- 427
QY 530 SYVRSRSIRSVAIEVDGRVYHVGGLDAAQPRNLTKRHWPAGAPDQDDKQDGFSGTGLP 589
Db 428 -----RNV----- 431
QY 590 DYSAANPIKVTHRCYILENTVQCDDLDLYKSLQAWKQHLHIDHEIETLQNKILQREVR 649
Db 432 ----- 431
QY 650 GHLKKRPEECCHKISVYTHQKGLKRGSSLLHPRKGLQKQKWLRLRQKKKLKR 709
Db 432 -----DPTCPSLS----- 439
QY 710 LLKGLQNDTCSMPGLT--CFTHDQHWQTAFTWTLGPFCACTSANNTYWCMTINETHN 768
Db 440 -----PQVSOCP-----PDCVCEADYNNYACVRTMSELWN 470
QY 769 FLFCFEE--ATGFELEYFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCYKQCNPR 825
Db 471 LOYCEFFDQEVFEVYVNLTAHPHOLNIAKSIDPELLGKNYRLMLMLQSCSGTCTPR 529

RESULT 3
ID - ARS VOLCA STANDARD; PRT; 649 AA.
AC Q10723;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arylsulfatase precursor (EC 3.1.1.6.1) (Aryl-sulfate sulphohydrolase).
```

```
OS Volvox carteri.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Volvocaceae; Volvox.
OX NCBI_TaxID=3067;
RN 1
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RX STRAIN=f. Nagariensis / HK10;
RX MEDLINE=94222055; PubMed=8168504;
RA Hallmann A., Sumper M.;
RT "An inducible arylsulfatase of Volvox carteri with properties
RT suitable for a reporter-gene system. Purification, characterization
RT and molecular cloning.";
RL Eur. J. Biochem. 221:143-150(1994).
RN 2
RP SEQUENCE OF 64-76, AND POST-TRANSLATIONAL MODIFICATION OF CYS-72.
RX MEDLINE=96283826; PubMed=8681943;
RA Selmer T., Hallmann A., Schmidt B., Sumper M., von Figura K.;
RT "The evolutionary conservation of a novel protein modification, the
RT conversion of cysteine to serinesemialdehyde in arylsulfatase from
RT Volvox carteri.";
RL Eur. J. Biochem. 238:341-345(1996).
CC -!- FUNCTION: Is commonly produced by soil microorganisms and plays an
CC important role in the mineralization of sulfates.
CC -!- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
CC -!- ENZYME REGULATION: INHIBITED BY NA(3)BO(3) AND KCN. NO INHIBITION
CC BY SODIUM DODECYL SULFATE, EVEN AT HIGH CONCENTRATION.
CC -!- SUBCELLULAR LOCATION: Periplasmic.
CC -!- INDUCTION: By sulfur deprivation.
CC -!- MISCELLANEOUS: The enzyme is thermostable, exhibiting a
CC temperature optimum at 60 degrees Celsius. Its optimal pH is 8.0.
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; X77214; CA54426.1; -.
DR PIR; S43229; S43229.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
DR PROSITE; PS00149; SULFATASE_2; 1.
KW Hydrolase; Signal; Glycoprotein; Periplasmic.
FT SIGNAL 1 22
FT CHAIN 23 649 ARYLSULFATASE.
FT MOD RES 72 72 2-AMINO-3-OXOPROPIONIC ACID.
FT DOMAIN 571 582 POLY-ALA.
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 224 224 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 649 AA; 72287 MW; 0C23BFD77C43F7B9 CRC64;

Query Match
Best Local Similarity 27.4%; Score 452.5; DB 1; Length 649;
Matches 149; Conservative 82; Mismatches 170; Indels 143; Gaps 24;

QY 6 LVLCILSATVSLGGSSAFSLSHRLKGRFDRNRIRPNILVLTDQDVELGSM--QV 63
Db 7 VALCLLG---FAALTAAAA---HQ-----RPNFVWIFTDDQDGIQNSTHPRY 47
QY 64 MNKTRIMEQGGTHFINAFVTPMCCPSRSSILTKYVHNHTYTNENCSSP-----SW 118
Db 48 QPKLHEHRYPGIELKNYFVTPVCCPSRTNLRGQFSHN-----TNFTDVLGPHGQYAKW 103
QY 119 QAQHSRTF-AVYLNSTGYRTAFPGKYLNEY---NGSYVPPGWEVGLLKNRFR-YYNT 173
```

```

Db      104 KSLGIDKSYLPVWLQNLQNTYVVGKFLVDYVSQYQNVNPGAGTIDIALVTPYTFDYNNP 163
QY      174 -LCRNGVKEKGSYKSDVLTDLTNDVSFRTSKMYPHRPVLMVLSHAAPHGPEDSA 232
Db      164 GFSRNGATP---NIVPGFSTVDIADKVAQIKTA--VAAGKPFYQAIPIAPH----- 212
QY      233 PQYSRLFPNASQHTPSYNAENP-DKIMIMRYTGPM-----KPIHMEF 275
Db      213 -TSTQIYFDPVANATKTFYPPIPAPRHWELESDATLPEGTSHKNLYEADVSDKPAMIRA 271
QY      276 TNNMQ-----RKELOTMSVDDSMETIYNMLVETGELDNTIYVYADHCYHIGQ 324
Db      272 LPLAQNNNTYLEEVYRLRSLASVDELIDRVATLQEAAGVLDNTIYIYSDNGSYHVT 331
QY      325 FGLVKGKMPYEFDIRVPYVRGPNVEAG-----CLNPHIVMLIDILAPTILDIAG----- 374
Db      332 HREGACKVTAYDEDLRVPFLIRGPGIRASHDKPANSKVGHLHVDFAPIILTLAGAGDQVG 391
QY      375 -----LDIPADMGKSIKLLDTERPVRPHLKXGVRWDSFLVE-----RGKLLH 421
Db      392 DKALDGTPLGLYANDGN---LLADYPRANH-----RNQFQGEFGWGSDEVLH 438
QY      422 KRNDKVDQAENEFLPKYQ-----RVKDLQRAEYOTACEOLGQKQCVEDATGKXL 474
Db      439 -----HIPRYTNSKAVRYD-----EDNQAKLIVSCTNEREL 474
QY      475 HKCK 478
Db      475 YDLK 478

```

RESULT 4

```

ARS CHLRE STANDARD; PRT; 647 AA.
ID ARS CHLRE STANDARD; PRT; 647 AA.
AC P14217;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase).
GN Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-55.
RC STRAIN=CW15;
RX MEDLINE=89384447; PubMed=2476654;
RA de Hostos E.L., Schilling J., Grossman A.R.;
RT "Structure and expression of the gene encoding the periplasmic
RT arylsulfatase of Chlamydomonas reinhardtii.";
RL Mol. Gen. Genet. 218:229-239 (1989).
RN [2]
RP IDENTIFICATION OF PROBABLE FRAMESHIFTS.
RA Bairoch A.;
RA Unpublished observations (AUG-1996).
RL -C- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
CC -I- SUBCELLULAR LOCATION: Periplasmic.
CC -I- INDUCTION: By sulfur deprivation.
CC -I- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC -I- CAUTION: THIS IS A CONCEPTUAL TRANSLATION, FRAMESHIFTS HAD TO BE
CC INTRODUCED TO MAXIMIZE THE SIMILARITY WITH OTHER SULFATASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X16180; CAA34302.1; ALT_FRAME.

```

```

DR EMBL; X16179; CAA34301.1; ALT_FRAME.
DR EMBL; X52304; CAA36545.1; -.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
DR PROSITE; PS00449; SULFATASE_2; 1.
KW Hydrolase; Signal; Glycoprotein; Periplasmic.
FT SIGNAL 1 21
FT CHAIN 22 647
FT MOD_RES 73 73
FT CARBOHYD 42 42
FT CARBOHYD 90 90
FT CARBOHYD 279 279
FT CARBOHYD 444 444
FT CARBOHYD 528 528
SQ SEQUENCE 647 AA; 72102 MW; 7404EA1BF233F0B1 CRC64;

Query Match 8.5%; Score 405; DB 1; Length 647;
Best Local Similarity 26.7%; Pred. No. 1.4e-21;
Matches 132; Conservative 76; Mismatches 179; Indels 108; Gaps 20;

QY 43 RNIILVLTDDQDELGSM--QVMNKRTRIMEQGGTHFINAFVITPMCPSSRSLSITKY 100
Db 26 KFNFFVIFTDDQAIQNSTHYPMLSHKYIRYPCGVLSQYFVTTVPCPSRTNLXRGQF 85
QY 101 VHNHTYTNENCSGP-----SWQAQHSRTF-AVYLNSTGYRTAFPGKYLNEY---NGS 151
Db 86 AHN-----TNFTSVLPFGWAKMGIGIDQSYLPLWLKDDQGYNTYVVGKFLVDYSVSYQ 141
QY 152 YVP-PGMKEWVGLLKNRSFYNTLCRNGVKENKHSYSDSKYLDLTLDITNDSVSFFTSKM 210
Db 142 QVPRAGTISMPXVTPYTFDYNTLQNGATP---NIYGEYSTVDVIRKGVQAIKSA--V 196
QY 211 YPHRPVLMVISHAAPHGPEDSAPOYSRLFPNASQHTPSYNAENP-DKIMIMRYTGPM- 268
Db 197 AGKPFYQAIPIAPH-----TSTQISTNPATGVTRSYFFPPIPAHPHMFSDANLP 249
QY 269 -----KPI-----HMEFTNMLQKRLQTLMSVDDSMETIYNMLV 302
Db 250 GGSXNKNLYEVDVSDKPAWIRALPLAQNNRTYQEEIYELRLRSL-GPELIEIQVVKTL 308
QY 303 ETGELDNTIYVYADHCYHIGQFGLVKGKMPYEFDIRVPYVRGPNVEAG---CLNPH 358
Db 309 EAGVLDNTIYIYSDNGYHVGARFGAGKTTGYEEDLRVPFLIRGPGIKASKSDKPNK 368
QY 359 IVNLIDILAPTILDIAG-----LDIPADMGKSIKLLDTERPVRPHLKXK 404
Db 369 VGLHVDPAFTILSLAGASHLLGDKLGDGTPLGLYANDG-----TLPEDYRPEHQROF 423
QY 405 MRV-----WRDSFLVERGKLLHKRNDKVDQAQENPLPKYQVRVKDLQRAEYOTACEQLGQ 460
Db 424 QGEFGWGSDELL-----QNLRSQPNNTW-KVRYTD-----ESSKQ 459
QY 461 KWQCVEDATGKXLH 475
Db 460 GWKLIQAQCTNERELY 474

RESULT 5
YIDJ ECOLI STANDARD; PRT; 497 AA.
ID YIDJ ECOLI STANDARD; PRT; 497 AA.
AC P31447;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative sulfatase YidJ (EC 3.1.6.-).
GN YIDJ OR B3678.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Bacterichia.
OX NCBI_TaxID=562;
RN [1]

```



```

Db 158 --ITWQMEYDDVAFANOKUQLSRENDDESRRPWCLTVSTFTHPHDPYVARRKFWDLYE 215
Qy 241 NASQHTSYNAP-----NPKKHLMRYTGPMKPIHMEFTNM-----LQKR--LQTL 287
Db 216 DC-EHLTPVGAIPLEQPHSQRM-----LSCDYQFQDTEVRNRSRYPANI 266
Qy 288 MSVDDSMETINMLVETGELDNTYIVYADHGHYQGLVKGKSMPEFDIRVFFVYRG 347
Db 267 SYLDKVKVGLIDTLRTLRMLDITLFLSCDHGDMGLGERGL-WFKMNFPEGSARVPLMIAG 325
Qy 348 PNVEAGC-LNPHIVLNIDLAPTILDIAGL--DIPADMKGKSLKLLD 391
Db 326 PGIAFGLHLP--TSNLDVFTPLADLAGISLEVRPWTGSLVPMWN 371

RESULT 7
ID YDEN ECOLI STANDARD; PRT; 560 AA.
AC P77318; P78159;
DT 16-OCT-2001 (Rel. 40, Created)
DE Putative sulfatase yden precursor (EC 3.1.6.-).
GN YDEN OR B1498.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R.; Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitagawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377 (1996).
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF000247; AAC74571.1; ALT INIT.
CC DDBL; D90791; BAA15169.1; ALT INIT.
CC EMBL; D90792; BAA15172.1; ALT_INIT.
CC EcoGene; EG13796; yden.
CC InterPro; IPR000917; Sulfatase.
CC Pfam; PF00884; Sulfatase; 1.
CC PROSITE; PS00523; SULFATASE_1; FALSE_NEG.
CC PROSITE; PS00149; SULFATASE_2; 1.
CC KW Hypothetical protein; Hydrolase; Signal; Complete proteome.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 560 PUTATIVE SULFATASE YDEN.
FT ACT_SITE 185 185 POTENTIAL.

```

```

SQ SEQUENCE 560 AA; 62802 MW; 67ED7FF7696C7A9F CRC64;
Query Match 5.3%; Score 251; DB 1; Length 560;
Best Local Similarity 22.7%; Pred. No. 1.6e-10;
Matches 129; Conservative 75; Mismatches 187; Indels 178; Gaps 25;

Qy 18 LLGGSASFUSHLKGRFORRRNI-----RPNILVLTDQ----- 54
Db 18 LASGMAFAAHAADDVKLKATKTNVAFSDFTPTYSTKGKNNIIVLTMDDLGCGQLPEDK 77
Qy 55 -----DVLGSMQVMNKTIRIMEOGGTHFINAFVYTPMCCPRS 93
Db 78 GGFDPKTMENREVDYIKIGIDKAEAAQKSTPTLLSLMDEGVRFNGYVAHGVSGPSRA 137
Qy 94 SILTKGYVHNHTYTNENCCSPSWQAQHE---SRTFAVYL-NSTGYRTAFPGK-YLNEY 148
Db 138 AIMGTRAPARFGVYSNTD-----AQDIGPLTETFLPELPQNHGYVYTAAGVKHLSKI 189
Qy 149 NGSYPPGKWEVGLLKNSRFY--NYTL--CRNGVKERHGSY----- 187
Db 190 SNVPVPED-----KQTRDYHDFNFTTFSAEEMQPNQRFDFMGFHAAGTAYYNSPSL 241
Qy 188 -----SKYLTDLITNDSVSFFRTSKMYPHPRLVLMVISHAAPHGPEDS-AP-QYSR 237
Db 242 FKXREVPKAGYISDQIDAEIGVDRAKTL--DQPFMLYLAYNAPHLPNDNPADQYOK 299
Qy 238 LFNASQHIPTSYNAPNDKHMRYTGPMPKPIHMEFTNMLQKRKLQTLMSVDDSMETI 297
Db 300 QENTGSQ--TADNYA-----SVYSVQCGVKRI 325
Qy 298 YNMLVETGELDNTYIVYADHGHYI-GQF---GLVKG-KSMPEYEDIRVPFVYRGNVEA 352
Db 326 LSQLKXNGYDNTIILFTSDNGAVTDGFLPLNGAQKYSQTSYFGTHTPMPMMW----K 381
Qy 353 GCINP-----HIVLNIDLAPTILDIAGLDIPAD--MDGKSILKLLDTER---PVNRFHLKK 403
Db 382 GKLPQNYDKLISAMDFTALDAADISIPKDLKLDGVSLLPLWLDKKGQEPHKLWTIT 441
Qy 404 KMRVRDSFLVERGKLLHK-----RDND--KYDAQENFL 436
Db 442 SYSHMFDEENIPFDNYHKFVRHOSDDYPHNPNTEDLSQFSYTVRNNDSLVYTVNNQL 501
Qy 437 PKYQVKKLCQRAEYCTACEQLGQKQCV 465
Db 502 GUY-KLTDLQKDNLAANPQVVKQGV 529

RESULT 8
STS HUMAN
ID STS HUMAN STANDARD; PRT; 583 AA.
AC P08842;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Steryl-sulfatase precursor (EC 3.1.6.2) (Steroid sulfatase) (Steryl-
DE sulfatase) (Steryl-sulfatase) (Steryl-sulfatase) (Steryl-
GN STS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89340479; PubMed=2668275;
RA Stein C., Hille A., Seidel J., Rijnbout S., Waheed A., Schmidt B.,
RA Geuze H., von Figura K.;
RT "Cloning and expression of human steroid-sulfatase. Membrane
RT topology, glycosylation, and subcellular distribution in BHK-21
RT cells.";
RL J. Biol. Chem. 264:13865-13872 (1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87187642; PubMed=3032454;

```

RA Yen P.H., Allen E., Marsh B., Mohandas T., Wang N., Taggart R.T.,
 RA Shapiro L.J.;
 RT "Cloning and expression of steroid sulfatase cDNA and the frequent
 RT occurrence of deletions in STS deficiency: implications for X-Y
 RT interchange.";
 RL Cell 49:443-454 (1987).
 RN [3]
 RP SEQUENCE OF 134-174 AND 461-583 FROM N.A.
 RX MEDLINE=89077541; PubMed=3203382;
 RA Yen P.H., Marsh B., Allen E., Tsai S.P., Ellison J., Connolly L.,
 RA Neiswanger K., Shapiro L.J.;
 RT "The human X-linked steroid sulfatase gene and a Y-encoded
 RT pseudogene: evidence for an inversion of the Y chromosome during
 RT primate evolution.";
 RL Cell 55:1123-1135 (1988).
 RN [4]
 RP SEQUENCE OF 22-45.
 RC TIGSUE=Liver;
 RX MEDLINE=89352671; PubMed=2765556;
 RA Kawano J.-I., Kocani T., Ohtaki S., Minamino N., Matsuo H., Oinuma T.,
 RA Aikawa E.;
 RT "Characterization of rat and human steroid sulfatases.";
 RL Biochim. Biophys. Acta 997:199-205 (1989).
 RN [5]
 RP VARIANTS XLI LEU-341; ARG-372 AND TYR-446.
 RX MEDLINE=92170784; PubMed=1539590;
 RA Basier E., Grompe M., Parenti G., Yates J., Ballabio A.;
 RT "Identification of point mutations in the steroid sulfatase gene of
 RT ichthyosis. Effects on the structure and function of the steroid
 RT sulfatase protein.";
 RL J. Biol. Chem. 272:20756-20763 (1997).
 RN [7]
 RP VARIANTS XLI PRO-560.
 RX MEDLINE=20146224; PubMed=10679952;
 RA Sugawara T., Shimizu H., Hoshi N., Fujimoto Y., Nakajima A.,
 RA Fujimoto S.;
 RT "PCR diagnosis of X-linked ichthyosis: identification of a novel
 RT mutation (B560P) of the steroid sulfatase gene.";
 RL Hum. Mutat. 15:296-296 (2000).
 RN [8]
 RP VARIANTS XLI ARG-380.
 RX MEDLINE=20304877; PubMed=10844566;
 RA Oyama N., Satoh M., Iwatsuki K., Kaneko F.;
 RT "Novel point mutations in the steroid sulfatase gene in patients with
 RT X-linked ichthyosis: transfection analysis using the mutated genes.";
 RL J. Invest. Dermatol. 114:1195-1198 (2000).
 CC -!- FUNCTION: CONVERSION OF SULFATED STEROID PRECURSORS TO ESTROGENS
 CC DURING PREGNANCY.
 CC -!- CATALYTIC ACTIVITY: 3-beta-hydroxyandrost-5-en-17-one 3-sulfate +
 CC H(2O) = 3-beta-hydroxyandrost-5-en-17-one + sulfate.
 CC -!- SUBUNIT: Homodimer.
 CC -!- SUBCELLULAR LOCATION: MICROSMAL MEMBRANE. THE SEQUENCE SHOWS
 CC SEVERAL MEMBRANE-SPANNING DOMAINS THAT COULD SERVE TO ANCHOR THE
 CC PROTEIN IN THE MICROSMAL MEMBRANE.
 CC -!- DISEASE: Defects in STS are the cause of x-linked ichthyosis
 CC (XLI), a diskarization disorder characterized by the presence
 CC of prominent scales, especially on the neck, extremities, trunk,
 CC and buttocks. It affects approximately 1 of 2000-6000 males.
 CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL; J04964; AAA60597.1; -;
 DR EMBL; M16505; AAA60596.1; -;
 DR EMBL; M23945; AAA60598.1; -;
 DR EMBL; M23556; AAA60599.1; -;
 DR PIR; A32641; KJHUAC.
 DR HSSP; F15846; IFSU.
 DR Genew; HGNC:11425; STS.
 DR MIM; 308100; -;
 DR GO; GO:0005783; C:endoplasmic reticulum; TAS.
 DR GO; GO:0005768; C:endosome; TAS.
 DR GO; GO:0005794; C:Golgi apparatus; TAS.
 DR GO; GO:0005764; C:lysosome; TAS.
 DR GO; GO:0005624; C:membrane fraction; TAS.
 DR GO; GO:0005792; C:microsome; TAS.
 DR GO; GO:0005886; C:plasma membrane; TAS.
 DR GO; GO:0004773; F:steryl-sulfatase activity; TAS.
 DR GO; GO:0008544; P:epidermal differentiation; TAS.
 DR GO; GO:0006706; P:steroid catabolism; TAS.
 DR InterPro; IPR000917; Sulfatase.
 DR Pfam; PF00884; Sulfatase; 1.
 DR PROSITE; PS00523; SULFATASE 1; 1.
 DR PROSITE; PS00149; SULFATASE 2; 1.
 DR KW Hydrolase; Microsome; Transmembrane; Glycoprotein; Steroid metabolism;
 KW Pregnancy; Signal; Disease mutation.
 FT SIGNAL 1 21
 FT CHAIN 22 583
 FT MOD_RES 75 75
 FT
 FT ACT_SITE 136 136
 FT TRANSMEM 185 206
 FT TRANSMEM 213 234
 FT CARBOHYD 47 47
 FT CARBOHYD 259 259
 FT CARBOHYD 333 333
 FT CARBOHYD 459 459
 FT VARIANT 341 341
 FT
 FT VARIANT 372 372
 FT
 FT VARIANT 372 372
 FT
 FT VARIANT 380 380
 FT
 FT VARIANT 444 444
 FT
 FT VARIANT 446 446
 FT
 FT VARIANT 560 560
 FT
 FT CONFLICT 23 23 A -> E (IN REF. 2).
 FT SEQUENCE 583 AA; 65492 MW; 74746AFA9D21A0A6 CRC64;
 SQ
 Query Match 5.0%; Score 238.5; DB 1; Length 583;
 Best Local Similarity 20.7%; Pred. No. 1.3e-09;
 Matches 126; Conservative 67; Mismatches 180; Indels 237; Gaps 23;
 QY 43 RPNILVLDGDDVGLGSMQVM-NXTRRI-----MEOGTHFINAFVTPMCCPSRSIL 96
 DB 26 RPNILVMD--DLGIDPGCYGKNTIRTPNIDRLASGVKLTQHLAASPLCTPSRAAFM 83
 QY 97 TGYVHNHTYNNENCSPSMQ-----AQHESRTFAVLYNSTGYRTAFEGK 143
 DB 84 TGRY-----PVRSGMASWSRTGVFLFTASSGGLPTBITPAKLLKDDGYSTALIGK 134
 QY 144 Y-----LNEY-----NGSYVPPGWKE----- 159
 DB 135 WHLGMSCHSKTDFCHPLHGHGNYFYGISLTNLRDKPEGSGVFTTGKRLVFLPLQIVG 194
 QY 160 -----WVGLLKNRFRNYNTLCRNGVKRKHG 184
 DB 195 VTLLTLAALNCGLLHVPLGVFPLFTASSGGLPTBITPAKLLKDDGYSTALIGK 252

QY 185 SDVSKDYLDLTNDVSFRTSKMYPHRPVLVMSHAAPHGPDSPAPQSRILFPNASQ 244
 DB 253 QPMSTNLTQRLTVEAQQIQRNTE-----TPFLVLVSLVHV-----TALFSSKDFAGKSQ 304
 QY 245 HITPSYNYAPNPDKHWIMRYTQPMKPIHMEFTNMLQRLKQLTSLVSDSDMETIYNMLVET 304
 DB 305 H--GVYGD-----VEMDMVSQIILNLLDEL 329
 QY 305 GELDNTIYVYVYADHGVHI-----GQGLVK-GKSMPEYBPDIRVPFVVRGPNV-E 351
 DB 330 RLANDLTFTSDQGAHVEVSKGEIHGSGNGIYKGRANNWEGGIRVPGILLRPVIQ 389
 QY 352 AGCLNPHIVNLIDLAFTIIDLADIPAD--MDGKSILKLD--TERPVN----- 397
 DB 390 AGQKIDPTSNMDIFPTVAKLAGAPLPEBRIIDGRDLPLLEGKQSRSDHEFLPHYCNAY 449
 QY 398 ----RFLKKKVRVWDSFLVER-----GKLLHKRD-----NDKVIDA 430
 DB 450 LNAVRWHPONSTSIWKAFFFTFNPNVSGNGCFATHVCFGSGYVTHDPPLLFDISKDP 509
 QY 431 QEENFL-----PKYQKVCLQRAE-----YQFACEQLGQK 461
 DB 510 RERNPLTPASEPRFYELKVMQEAERDHTQTLPPEVDQFSMNFLWKPMQLQCCFSTGLS 569
 QY 462 WQCVEDATGK 471
 DB 570 CQCDREKQDK 579

RESULT 9
 ARS_STRPU STANDARD; PRT; 567 AA.
 AC P50473;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase) (ARS).
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinozoa; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89357267; PubMed=2767335;
 RA Yang Q., Angerer L.M., Angerer R.C.;
 RT "Structure and tissue-specific developmental expression of a sea urchin arylsulfatase gene";
 RL Dev. Biol. 135:53-65(1989).
 CC -!- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: EXPRESSION IS CONFINED TO ABORAL ECTODERM CELLS AND THEIR PRECURSORS.
 CC -!- DEVELOPMENTAL STAGE: LOW LEVELS ARE FOUND AT MESENCHYME BLASTULA STAGE (24 HR), LEVELS INCREASE BY LATE GASTRULA STAGE AND ARE MAINTAINED AT PLUTEUS STAGE.
 CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC EMBL; M28404; AAA30036.1; .
 CC PIR; A37362; A37362.
 CC HSSP; P15289; 1AUK.
 CC InterPro; IPR000917; Sulfatase.
 CC Pfam; PF00884; Sulfatase; 1.

DR PROSITE; PS00523; SULFATASE 1; 1.
 DR PROSITE; PS00149; SULFATASE 2; 1.
 KW Hydrolase; Signal; Glycoprotein.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 567 ARYL SULFATASE.
 FT MOD_RES 115 115 2-AMINO-3-OXOPROPIONIC ACID (BY SIMILARITY).
 FT ACT_SITE 173 173 POTENTIAL.
 FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 567 AA; 62477 MW; D5B627983A4C4D6 CRC64;

Query Match 5.0%; Score 238; DB 1; Length 567;
 Best Local Similarity 23.4%; Pred. No. 1.4e-09;
 Matches 103; Conservative 58; Mismatches 167; Indels 112; Gaps 18;

QY 2 GPSSLVLCILSATVFLSGSSAFSLSHRLKGRFORDRENIRPNILVLTDDQDVGLGSM 61
 DB 27 GPDABSLASLDRTATRRYGDGEDLL--HLGGTQGHRTAMTKPNVILLADDMGV--GDL 82
 QY 62 QVANKTRR-----IMEQGTGFHINAFVTTMCCPSRSSILTKYVHNHTYTNNE---- 111
 DB 83 SVYGHPTQBPFGFIDQMANQGLRFTQGYSGDSVCTPSGSAIVTGRQPIRTGVYGEERIFLP 142
 QY 112 --NCSSPSWQAQHSRTFAVYLNSTGYRTAPFGK--LNE---YNGSYVP----- 154
 DB 143 WTTTGLPLYEV-----TIAEAMKAGYTTGMVGMKHLGINENSSDGAHLPANRGDFVFG 197
 QY 155 ---PGWKBN---VGLLKN-----SRFYNTLCRNGVKKEKHSYDSDYLDLTITNDS 200
 DB 198 HNLPGNSWRCDTGLHQDPDTNACFLYNTSVTAQPFQHKG-----LTQLLRDRT 249
 QY 201 VSFRTSKMYPHRPVLVMSHAAPHGPDSPAPQSRILFPNASQHIPTSYNAPNPKHW 260
 DB 250 VGFIEDN---VNKPFMYVSPAHME-----TSLFSSD----- 278
 QY 261 IMEYTCGPMKPIHMEFTNMLQRLKEL--QTLSVDDSDMETIYNMLVETGEIDNTYIVVYADHG 319
 DB 279 -----DFCTSRGRYGDNLREMQALIEQIVTTLVDNDIDNTVFTISHG 325
 QY 320 YH---IQGFGLVK---GKSMPEYBPDIRVPFVVRGPNVEAGCLNPHIVNLIDLAFTILDI 372
 DB 326 PHREYCGEGGDANVFRGKGQSWEGGHRIPYIVYVPGTISPGVSHEIVTSMIDIATAVNL 385
 QY 373 AGLDIPAD--MDGKSILKLL 390
 DB 386 GGSQLPTRIDYDGKCLKSVL 405

RESULT 10
 IDS_HUMAN STANDARD; PRT; 550 AA.
 ID IDS_HUMAN
 AC P22304; Q14604;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Iduronate 2-sulfatase precursor (EC 3.1.6.13).
 GN IDS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 34-58 AND 456-473.
 RC TISSUE=Endothelial cells;
 RX MEDLINE=91046030; PubMed=2122463;
 RA Wilson P.J., Morris C.P., Anson D.S., Occhiodoro T., Bielicki J.,
 RA Clements P.R., Hopwood J.J.;
 RT "Hunter syndrome: isolation of an iduronate-2-sulfatase cDNA clone and analysis of patient DNA";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8531-8535(1990).
 RN [2]

RP SEQUENCE FROM N.A.
RX MEDLINE=94063929; PubMed=8244397;
RA Wilson P.J., Meaney C.A., Hopwood J.J., Morris C.P.;
RT "Sequence of the human iduronate-2-sulfatase (IDS) gene.";
RL Genomics 17:773-775(1993).
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=96352905; PubMed=8717057;
RX Timms K.M., Lu F., Shen Y., Pierson C.A., Muzny D.M., Gu Y.,
RA Nelson D.L., Gibbs R.A.;
RT "130 kb of DNA sequence reveals two new genes and a regional
RT duplication distal to the human iduronate-2-sulfate sulfatase
RT locus";
RL Genome Res. 5:71-78(1995).
[4]
RN SEQUENCE FROM N.A.
RP MEDLINE=94154729; PubMed=8111411;
RX Hopwood J.J., Bunge S., Morris C.P., Wilson P.J., Steglich C.,
RA Beck M., Schwinger E., Gal A.;
RT "Molecular basis of mucopolysaccharidosis type II: mutations in the
RT iduronate-2-sulfatase gene.";
RL Hum. Mutat. 2:435-442(1993).
[8]
RN VARIANTS MPS-II ARG-135 AND GLY-422.
RX MEDLINE=93265059; PubMed=1303211;
RA Bunge S., Steglich C., Beck M., Rosenkranz W., Schwinger E.,
RA Hopwood J.J., Gal A.;
RT "Mutation analysis of the iduronate-2-sulfatase gene in patients with
RT mucopolysaccharidosis type II (Hunter syndrome).";
RL Hum. Mol. Genet. 1:335-339(1992).
[9]
RN VARIANTS MPS-II TRP-468.
RX MEDLINE=93258320; PubMed=1284597;
RA Crotti P.L., Bunge S., Anderson R.A., Whitley C.B.;
RT "Mutation R468W of the iduronate-2-sulfatase gene in mild Hunter
RT syndrome (mucopolysaccharidosis type II) confirmed by in vitro
RT mutagenesis and expression";
RL Hum. Mol. Genet. 1:755-757(1992).
[10]
RN VARIANTS MPS-II ARG-86; ASP-94; ARG-120; PRO-221 AND GLY-422.
RX MEDLINE=94108441; PubMed=8281149;
RA Bunge S., Steglich C., Zuther C., Beck M., Morris C.P., Schwinger E.,
RA Schinzel A., Hopwood J.J., Gal A.;
RT "Iduronate-2-sulfatase gene mutations in 16 patients with
RT mucopolysaccharidosis type II (Hunter syndrome).";
RL Hum. Mol. Genet. 2:1871-1875(1993).
[11]
RN VARIANTS MPS-II GLU-68; HIS-293; GLY-478 AND ARG-485.
RX MEDLINE=95072615; PubMed=7981716;
RA Schroeder W., Wulff K., Wennert M., Seidlitz G., Herrmann F.H.;
RT "Mutations of the iduronate-2-sulfatase (IDS) gene in patients with
RT Hunter syndrome (mucopolysaccharidosis II).";
RL Hum. Mutat. 4:128-131(1994).
[12]
RN VARIANTS MPS-II TRP-132; TYR-229; ARG-358; HIS-469 AND CYS-523.
RX MEDLINE=95193786; PubMed=7887413;
RA Jonsson J.J., Aronovich E.L., Braun S.E., Whitley C.B.;
RT "Molecular diagnosis of mucopolysaccharidosis type II (Hunter
RT syndrome) by automated sequencing and computer-assisted
RT interpretation: toward mutation mapping of the iduronate-2-sulfatase
RT gene.";
RL Am. J. Hum. Genet. 56:597-607(1995).
[13]
RN VARIANTS MPS-II LEU-86; ASN-87; PRO-92; CYS-345 AND TRP-468.
RX MEDLINE=95245347; PubMed=7728156;
RA Popowska E., Rathmann M., Tytki-Szymanska A., Bunge S., Steglich C.,
RA Schwinger E., Gal A.;
RT "Mutations of the iduronate-2-sulfatase gene in 12 Polish patients
RT with mucopolysaccharidosis type II (Hunter syndrome).";
RL Hum. Mutat. 5:97-100(1995).
[14]
RN VARIANTS MPS-II VAL-346.
RX MEDLINE=95322987; PubMed=7599640;
RA Li P., Huffman P., Thompson J.N.;
RT "Mutations of the iduronate-2-sulfatase gene on a T146T background in
RT three patients with Hunter syndrome.";
RL Hum. Mutat. 5:272-274(1995).
[15]
RN VARIANTS MPS-II.
RX MEDLINE=97094177; PubMed=8940265;
RA Rathmann M., Bunge S., Beck M., Kresse H., Tytki-Szymanska A., Gal A.;
RT "Mucopolysaccharidosis type II (Hunter syndrome): mutation 'hot spots'
RT in the iduronate-2-sulfatase gene.";
RL Am. J. Hum. Genet. 59:1202-1209(1996).
[16]
RN VARIANTS MPS-II LEU-333 AND ASP-346.
RX MEDLINE=96163494; PubMed=8566953;
RA Olsen T.C., Eiken H.G., Knappskog P.M., Kase B.F., Mansson J.-E.,
RA Roman H., Apold J.;
RT "Mutations in the iduronate-2-sulfatase gene in five Norwegians with
RT Hunter syndrome.";
RL Hum. Genet. 97:198-203(1996).
[17]
RN VARIANTS MPS-II LEU-333 AND GLY-334.
RX MEDLINE=96255081; PubMed=8830188;
RA Li P., Thompson J.N.;
RT "Detection of four novel mutations in the iduronate-2-sulfatase gene
RT by single-strand conformation polymorphism analysis of genomic
RT amplicons.";
RL J. Inherit. Metab. Dis. 19:93-94(1996).
[18]
RN VARIANTS MPS-II ASP-63; THR-347; GLN-468 AND LEU-468.
RX MEDLINE=97365936; PubMed=9222763;
RA Villani G.R.D., Balzano N., Grosso M., Salvatore F., Izzo P.,
RA di Natale P.;
RT "Mucopolysaccharidosis type II: identification of six novel mutations
RT in Italian patients.";
RL Hum. Mutat. 10:71-75(1997).
[19]
RN VARIANTS MPS-II GLN-468.
RX MEDLINE=98041699; PubMed=9375851;
RA Sukegawa K., Song X.-Q., Masuno M., Fukao T., Shimozaawa N., Fukuda S.,
RA Isozaki K., Nishio H., Matsuo M., Tomatsu S., Kondo N., Orii T.;
RT "Hunter disease in a girl caused by R468Q mutation in the
RT iduronate-2-sulfatase gene and skewed inactivation of the X
RT chromosome carrying the normal allele.";
RL Hum. Mutat. 10:361-367(1997).
[20]
RN VARIANTS MPS-II ASN-45; TYR-115; LEU-228; ARG-266; LYS-434; LYS-485
RP AND CYS-502.
RX MEDLINE=99092178; PubMed=9875019;
RA Vafiadaki E., Cooper A., Heptinstall L.E., Hatton C.E., Thornley M.,
RA Wraith J.E.;

"Mutation analysis in 57 unrelated patients with MPS II.";
Arch. Dis. Child. 79:237-241(1998).
[21]
RT VARIANTS MPS-II.
RA MEDLINE=9911886; PubMed=9921913;
RA Karsten S., Voskobojeva E., Tishkanina S., Pettersson U.,
RA Kraenopol'skaja X., Bondeson M.-L.;
RT "Mutational spectrum of the iduronate-2-sulfatase (IDS) gene in 36
RT unrelated Russian MPS II patients.";
RL Hum. Genet. 103:732-735(1998).
[22]
RA VARIANTS MPS-II LEU-86; HIS-88; PRO-88; ILE-118 AND HIS-266.
RA Balzano N., Villani G.R.D., Grosso M., Izzo P., di Natale P.;
RT "Detection of four novel mutations in the iduronate-2-sulfatase
RT gene.";
RL Hum. Mutat. 11:333-333(1998).
[23]
RA VARIANTS MPS-II THR-85; HIS-88; ILE-349 AND VAL-521.
RA MEDLINE=98112423; PubMed=9452044;
RA Gort L., Coll M.J., Chabas A.;
RT "Mutations in the iduronate-2-sulfatase gene in 12 Spanish patients
RT with Hunter disease.";
RL Hum. Mutat. Suppl. 1:S66-S68(1998).
[24]
RA VARIANTS MPS-II PHE-143; TRP-184; VAL-269 AND HIS-348.
RA Karsten S.L., Voskobojeva E., Carlberg B.-M., Kleijer W.J.,
RA Toennesen T., Pettersson U., Bondeson M.-L.;
RT "Identification of 9 novel gene mutations in 19 unrelated Hunter
RT syndrome (Mucopolysaccharidosis type II) patients.";
RL Hum. Mutat. 12:433-433(1998).
[25]
RA VARIANTS MPS-II PHE-143; LYS-341; TYR-342 AND PHE-491.
RA MEDLINE=99235558; PubMed=10220152;

Query Match 4.9%; Score 234.5; DB 1; Length 550;
Best Local Similarity 21.2%; Pred. No. 2.4e-09;
Matches 114; Conservative 80; Mismatches 189; Indels 155; Gaps 23;

QY 1 MCPPS-----LVLCILLSATVFLSGSSAFSLSHRLKGRFQDRNRNIRPNILVLTDDQD 55
DB 1 MPPPTGRGLLGLVLSVVCVALGSET-----QANSTTDALNVLIIIVDDL 48
QY 56 VELGSMQVNNKTRIMEQOQTH---FINAVTTPMCCPSRSILTGK-----YVHNHTY 107
DB 49 PSLGCVGDKLVSPNIDQLASHSLFLQNAFAQAVCAPSRVSLTGRPDITRLYDNFSY 108
QY 108 TNENCSPSWAQAHES-RTFAVYLNSTGYRTAFPKYL-----NEYNGSVYP- 154
DB 109 -----MRVHAGNEFTIPQYFKNGYVTMSVGRVFEHGIGSNHNTDDSPYSNFPFY 158
QY 155 -PGWKEWGLLKNRFRNYTLCRNGYKEKHGS-----DYSKDYLD-LITNDSVSFF 204
DB 159 HP-----SSEKVENTKTCRGDGLHANLLCPVDVLPDPEGTLPDKOSTEQAIQLL 209
QY 205 RTSKMYPHRVLWVISHAHPGSDPAQYSRKLPFNASQHTIPSYNAPND----- 257
DB 210 ERKMT--SASPFLVAGYHKPHIPRYPKFQKLYP--LENIT---LAPDPEVPDGLPP 261
QY 258 ---KHWI-----MRYTGMPKPIHEMTNMLQRKELQTLMSVDDSMETIYNMLVET 304
DB 262 VAYNPMMDIQRQEDVOALNINISYPYGPVDFVDFORKIQSYFASVYLDTVQVRLSALDDL 321
QY 305 GELDNTYIVYTADGHYGHQGLVNGKSPYFDFIRVP--FYVRGN----- 349
DB 322 QLANSTIIAFTSDHGWALGEGH-EWAKYSNFDVATHVPLIFVFGRTASLPEAGEKLFY 380
QY 350 -----VRAGCLNPHIVLNIDLAFTILDIDIPADMKGKSIKLKLDTERPVNRF 399
DB 381 LOPFDSASQLMEPGRQMDLVLSLFTPLAGLAGLQVP-----PRCPVPSF 427
QY 400 HLKCKRVRVDFSLVERGKLLKHKRDNKVDYDAEENFLP-----KYGRVKDLQ 447
DB 428 HVE-----LCREGKRLKGFRR--DLBEDPYLPGNPRELIAYSOYRPSDIPQ 474

RESULT 11
ARS HEMPU STANDARD; PRT; 551 AA.
AC P1400; 1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Arylsulfatase precursor (BC 3.1.6.1) (Aryl-sulfate sulphohydrolase)
DE (ARS).
OS Hemacentrotus pulcherrimus (sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoida; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Hemacentrotus.
OX NCBI_TaxID=7650;
[1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP TISSUE=Pluteus;
RC MEDLINE=89030659; PubMed=3181160;
RX Sasaki H., Yamada K., Akasaka H., Suzuki K., Saito A., Sato M.,
RA Shimada H.;
RT "cDNA cloning, nucleotide sequence and expression of the gene for
RT arylsulfatase in the sea urchin (Hemacentrotus pulcherrimus)
RT embryo".
RL Eur. J. Biochem. 177:9-13(1988).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=90092130; PubMed=2598936;
RA Yamada K., Akasaka K., Shimada H.;
RT "Structure of sea-urchin arylsulfatase gene.";
RL Eur. J. Biochem. 186:405-410(1989).
CC -!- FUNCTION: MAY BE A STRUCTURAL COMPONENT OF THE EXTRACELLULAR
CC MATRICES INVOLVED IN CELL MOVEMENT DURING MORPHOGENESIS.
CC -!- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
CC -!- SUBCELLULAR LOCATION: IN BOTH THE CYTOPLASM AND THE EXTRACELLULAR
CC MATRICES.
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL; X17015; CAA34881.1; -.
DR PIR; S01793; S01793.
DR HSPP; P15289; LAUK.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
DR PROSITE; PS00149; SULFATASE_2; 1.
KW Hydrolase; Signal; Glycoprotein; Extracellular matrix.
FT SIGNAL 1 20
FT CHAIN 21 551 ARYLSULFATASE.
FT MOD_RES 21 21 BLOCKED.
FT MOD_RES 100 100 2-AMINO-3-OXOPROPIONIC ACID (BY
FT SIMILARITY).
FT ACT_SITE 158 158 POTENTIAL.
FT CARBOHYD 164 164 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 213 213 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 551 AA; 60952 MW; 54C1AAC14D6710C9 CRC64;

Query Match 4.9%; Score 232; DB 1; Length 551;
Best Local Similarity 23.6%; Pred. No. 3.7e-09;
Matches 105; Conservative 63; Mismatches 161; Indels 116; Gaps 21;
QY 42 IRPNILVLTDDQDVELGSMQVNNKTRIMEQO-----GTHFINAVTTPMCCPSRS 93
DB 50 VKPNVLLVAD-----HMGSGDLTSYGHPTQEGAFIDKMAAELGRFTNGVGVDAVCTPSRS 105

DR	PROSITE; PS00149; SULFATASE_2; 1.
KW	Hydrolase; Glycoprotein; Lysozyme; Zymogen; Signal.
SIGNAL	1 ? POTENTIAL.
FT	PROPEP ? 59 BY SIMILARITY.
FT	CHAIN 60 563 IDUONATE 2-SULFATASE.
FT	MOD_RES 99 2-AMINO-3-OXOPROPIONIC ACID (BY SIMILARITY).
FT	CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 526 526 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 550 550 N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQ	SEQUENCE 563 AA; 63437 MW; 52D008F91BA88FAC CRC64;
Query Match	4.8%; Score 226; DB 1; Length 563;
Best Local Similarity	21.7%; Pred. No. 1e-08;
Matches 101; Conservative	69; Mismatches 199; Indels 96; Gaps 17;
QY	16 FSLGGSSAFLSHRLKGRFQRDRNIRPNIILVTDODVELGSMQVMKTRIMEQGG 75
DB	28 FSLGGSFCTALESAAQGSNTDAL-----NILLIIVDLRSLCYGDKLVRSPIIDQLA 83
QY	76 TH---FINAFVTTPMCCPSRSILTKG-----YVNHNTYTNNECSPGQAOHES-RT 126
DB	84 SHSLFQNAFAQAQAVCAPSRVSLTGRRPDITRLDYDNSY-----WRVHSGNFST 133
QY	127 PAVYLNSTGVRTAFPKYL-----NEVNGSVYPPCKWEKVLKNSRFNYVTLCRNGVKE 181
DB	134 IPQYFKENGVTVMGSKGVKVFHPIGSISSNHSDDY-PYSNWSPPPHPSSEKYENTKTKGQDGK 192
QY	182 KHG-----SDYSKDYLTDLITNDSVFRTSKMYPH-RPVLMIWISHAAPHGPDSSA 232
DB	193 LHANLLCPVDVADVPSTLPDQSTEEA--IRLEKMTSGSPFLVAGVYKHPIPPRYP 250
QY	233 PQYRSLFPNASHITPSYVAPNDKH-----WI-----MRYTGMPKP 270
DB	251 KEFOKLYP--LENITILA-----PDHPVDSLPPVAYPNMWDIRERDVOALNISVYPG 302
QY	271 IHMEFTNMLORKELOTLMSVDDSMETIYNMLVETGELONTYIVTADHGVIHQFG--- 326
DB	303 IPEDFORKIRQSYFASVYLDITQVGHVLSALDLDLAENTIIAFTSDHGVALGHEGNAK 362
QY	327 -----LYKGSMPE-----PDIRVPYVPGNVEAGCLNPHIVLNTDLA 366
DB	363 YSNEDVATRVPLMLYVPGRTAPLPAAGOKLFPYRDPDPASDMWDAGRHTEDLVLSLF 422
QY	367 PTILDIAGLDIPADMDGKSTLKLLDTERPVNRFHLKKKRVWRDS 411
DB	423 PTLAGLAGL--PVLGAPSLPMLSPFAEKARIFRSCSMTWKS 465

RESULT 13

GA6S HUMAN

ID GA6S HUMAN STANDARD; PRT; 522 AA.

AC P34059.

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE N-acetylgalactosamine-6-sulfatase precursor (EC 3.1.1.6.4) (N-acetylgalactosamine-6-sulfate sulfatase) (Galactose-6-sulfate sulfatase) (GALNA6S sulfatase) (Chondroitin sulfatase).

DE GN GALNS.

OS Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE=Placenta;

RX MEDLINE=92095973; PubMed=1755850;

RA Tomatsu S., Fukuda S., Masue M., Sukeyasu K., Fukao T., Yamagishi A.,

RA Hori T., Iwata H., Ogawa T., Nakashima Y., Hanyu Y., Hashimoto T.,
RA Titani K., Oyama R., Suzuki M., Yagi K., Hayashi Y., Orii T.,
RT "Morquio disease: isolation, characterization and expression of full-
RT length cDNA for human N-acetylgalactosamine-6-sulfate sulfatase.";
RL Biochem. Biophys. Res. Commun. 181:677-683(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95095267; PubMed=8001980;
RA Morris C.P., Guo X.H., Apostolou S., Hopwood J.J., Scott H.S.;
RT "Morquio A syndrome: cloning, sequence, and structure of the human N-
RT acetylgalactosamine 6-sulfatase (GALNS) gene.";
RL Hum. Mutat. 22:652-654(1994).
RN [3]
RP VARIANTS MPS-IVA.
RX MEDLINE=92299122; PubMed=1522213;
RA Fukuda S., Tomatsu S., Masue M., Sukegawa K., Iwata H., Ogawa T.,
RA Nakashima Y., Hori T., Yamagishi A., Hanyu Y., Morooka K., Kiman T.,
RA Hashimoto T., Orii T.;
RT "Mucopolysaccharidosis type IVA. N-acetylgalactosamine-6-sulfate
RT sulfatase exonic point mutations in classical Morquio and mild
RT cases.";
RL J. Clin. Invest. 90:1049-1053(1992).
RN [4]
RP VARIANTS MPS-IVA.
RX MEDLINE=95397840; PubMed=7668283;
RA Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Maruf Rezvi G.,
RA Yamagishi A., Yamada N., Kato Z., Isogai K., Sukegawa K., Kondo N.,
RA Suzuki Y., Shimoza N., Orii T.;
RT "Mucopolysaccharidosis IVA: identification of a common missense
RT mutation I113F in the N-Acetylgalactosamine-6-sulfate sulfatase
RT gene.";
RL Am. J. Hum. Genet. 57:556-563(1995).
RN [5]
RP VARIANTS MPS-IVA.
RX MEDLINE=95315929; PubMed=7795586;
RA Ogawa T., Tomatsu S., Fukuda S., Yamagishi A., Maruf Rezvi G.,
RA Sukegawa K., Kondo N., Suzuki Y., Shimoza N., Orii T.;
RT "Mucopolysaccharidosis IVA: screening and identification of mutations
RT of the N-acetylgalactosamine-6-sulfate sulfatase gene.";
RL Hum. Mol. Genet. 4:341-349(1995).
RN [6]
RP VARIANTS MPS-IVA ARG-77; TRP-90; VAL-96; LEU-151; GLY-230 AND THR-291.
RX MEDLINE=95359983; PubMed=7633425;
RA Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Maruf Rezvi G.,
RA Yamagishi A., Yamada N., Kato Z., Isogai K., Sukegawa K., Kondo N.,
RA Suzuki Y., Shimoza N., Orii T.;
RT "Mucopolysaccharidosis type IVA: identification of six novel
RT mutations among non-Japanese patients.";
RL Hum. Mol. Genet. 4:741-743(1995).
RN [7]
RP VARIANTS MPS-IVA SER-487.
RX MEDLINE=96047158; PubMed=7581409;
RA Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Yamada N., Isogai K.,
RA Kato Z., Sukegawa K., Kondo N., Suzuki Y., Shimoza N., Orii T.;
RT "Two new mutations, Q473X and N487S, in a Caucasian patient with
RT mucopolysaccharidosis IVA (Morquio disease).";
RL Hum. Mutat. 6:195-196(1995).
RN [8]
RP VARIANTS MPS-IVA ALA-138; SER-151 AND LEU-151.
RX MEDLINE=96216146; PubMed=8631279;
RA Tomatsu S., Fukuda S., Yamagishi A., Cooper A., Wraith J.E., Hori T.,
RA Kato Z., Yamada N., Isogai K., Sukegawa K., Kondo N., Suzuki Y.,
RA Shimoza N., Orii T.;
RT "Mucopolysaccharidosis IVA: four new exonic mutations in patients
RT with N-acetylgalactosamine-6-sulfate sulfatase deficiency.";
RL Am. J. Hum. Genet. 58:950-962(1996).
RN [9]
RP VARIANTS MPS-IVA CYS-94 AND VAL-97.
RX MEDLINE=96423834; PubMed=8826435;
RA Cole D.E.C., Fukuda S., Gordon B.A., Rip J.W., Lecouteur A.N.,
RA Harper C.A., Tomatsu S., Ogawa T., Sukegawa K., Orii T.;
RT "Heteroallelic missense mutations of the galactosamine-6-sulfate
RT sulfatase (GALNS) gene in a mild form of Morquio disease (MPS IVA).";

Am. J. Med. Genet. 63:558-565(1996).
[10]
VARIANTS MPS-IVA.
MEDLINE=97442274; PubMed=9298823;
Bunge S., Kleijer W.J., Tytki-Szymanska A., Steglich C., Beck M.,
Tomatsu S., Fukuda S., Poorthuis B.J.H.M., Czartoryska B., Orii T.,
Gal A.;
"Identification of 31 novel mutations in the N-acetylgalactosamine-6-
sulfatase gene reveals excessive allelic heterogeneity among patients
with Morquio A syndrome.";
Hum. Mutat. 10:223-232(1997).
[11]
VARIANTS MPS-IVA.
MEDLINE=98041700; PubMed=9375852;
Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Ferreira P.,
di Natale P., Tortora P., Fujimoto A., Kato Z., Yamada N., Isogai K.,
Yamagishi A., Sukegawa K., Suzuki Y., Shimoza N., Kondo N.,
Sly W.S., Orii T.;
"Fourteen novel mucopolysaccharidosis IVA producing mutations in
GALNS gene.";
Hum. Mutat. 10:368-375(1997).
[12]
VARIANTS MPS-IVA.
MEDLINE=98180718; PubMed=9521421;
Yamada N., Fukuda S., Tomatsu S., Muller V., Hopwood J.J., Nelson J.,
Kato Z., Yamagishi A., Sukegawa K., Kondo N., Orii T.;
"Molecular heterogeneity in mucopolysaccharidosis IVA in Australia
RT and Northern Ireland: nine novel mutations including T112S, a common
RT allele that confers a mild phenotype.";
RL Hum. Mutat. 11:202-208(1998).
RN [13]
VARIANTS SER-393 AND MET-488.
MEDLINE=98112415; PubMed=9452036;
Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Yamagishi A., Kato Z.,
Yamada N., Isogai K., Sukegawa K., Suzuki Y., Shimoza N., Kondo N.,
Orii T.;
"Fifteen polymorphisms in the N-acetylgalactosamine-6-sulfate
RT sulfatase (GALNS) gene: diagnostic implications in Morquio disease.";
RL Hum. Mutat. Suppl. 1:S42-S46(1998).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 6-sulfate groups of the N-
CC acetyl-D-galactosamine 6-sulfate units of chondroitin sulfate and
CC of the D-galactose 6-sulfate units of keratan sulfate.
CC -!- SUBUNIT: OLIGOMER OF DISULFIDE LINKED 40- AND 15-KDA POLYPEPTIDES.
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- DISEASE: Defects in GALNS are the cause of mucopolysaccharidosis
CC type IVA (MPS-IVA) [MIM:253000]; also known as Morquio A syndrome.
CC MPS-IVA is characterized by specific spondyloepiphyseal dysplasia,
CC short trunk dwarfism, coxa valga, odontoid hypoplasia, corneal
CC opacities, preservation of intelligence, and excessive urinary
CC excretion of keratan sulfate and chondroitin-6-sulfate. Severely
CC affected patients usually die of cardiopulmonary disturbance or
CC cervical cord compression in the second or third decade of life.
CC -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D17629; BAA04535.1;
CC EMBL; D17616; BAA04535.1; JOINED.
CC EMBL; D17617; BAA04535.1; JOINED.
CC EMBL; D17618; BAA04535.1; JOINED.
CC EMBL; D17619; BAA04535.1; JOINED.
CC EMBL; D17620; BAA04535.1; JOINED.
CC EMBL; D17621; BAA04535.1; JOINED.
CC EMBL; D17622; BAA04535.1; JOINED.
CC EMBL; D17623; BAA04535.1; JOINED.
CC EMBL; D17624; BAA04535.1; JOINED.
CC EMBL; D17625; BAA04535.1; JOINED.

01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Arylsulfatase A precursor (EC 3.1.6.8) (ASA) (Cerebroside-sulfatase).
ARSA OR AS2.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
STRAIN=129/SV, and C57BL/6J;
MEDLINE=94245194; PubMed=7910580;
Kreysing J., Polten A., Hess B., von Figura K., Menz K., Steiner F.,
Gieselmann V.;
"Structure of the mouse arylsulfatase A gene and cDNA.";
Genomics 19:249-255(1994).
[2]
SEQUENCE OF 32-66 FROM N.A.
MEDLINE=92241876; PubMed=1572648;
Grompe M., Pieretti M., Caskey C.T., Ballabio A.;
"The sulfatase gene family: cross-species PCR cloning using the MOPAC
technique.";
Genomics 12:755-760(1992).
-!- FUNCTION: Hydrolyzes cerebroside sulfate.
-!- CATALYTIC ACTIVITY: A cerebroside 3-sulfate + H(2)O = a
cerebroside + sulfate.
-!- SUBUNIT: EXISTS BOTH AS A SINGLE CHAIN OF 58 Kda (COMPONENT A)
OR AS A CHAIN OF 50 Kda (COMPONENT B) LINKED BY DISULFIDE BOND(S)
TO A 7 Kda CHAIN (COMPONENT C).
-!- SUBCELLULAR LOCATION: Lysosomal.
-!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X73231; CAA51703.1; -;
EMBL; X73230; CAA51702.1; -;
EMBL; M82876; AAA37260.1; -;
PIR; A54190; A54190.
HSSP; P15289; LAUK.
MGD; MGI:88077; Arsa.
GO; GO:0000299; C:integral membrane protein of membrane fraction; IDA.
GO; GO:0005886; C:plasma membrane; IDA.
GO; GO:0007339; P:binding of sperm to zona pellucida; IMP.
InterPro; IPR000917; Sulfatase.
Pfam; PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
Hydrolase; Signal; Glycoprotein; Lysosome.
SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 506
FT MOD_RES 68 68
FT 2-AMINO-3-OXOPROPIONIC ACID (BY
SIMILARITY).
FT DISULFID 155 171
FT BY SIMILARITY.
FT DISULFID 160 167
FT BY SIMILARITY.
FT DISULFID 299 413
FT BY SIMILARITY.
FT DISULFID 487 499
FT BY SIMILARITY.
FT DISULFID 488 501
FT BY SIMILARITY.
FT DISULFID 492 498
FT BY SIMILARITY.
FT CARBOHYD 157 157
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 349 349
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQUENCE 506 AA; 53776 MW; 74BB16401CF18DF6 CRC64;

Query Match 4.4%; Score 210.5; DB 1; Length 506;
Best Local Similarity 25.2%; Pred. No. 1.2e-07;
Matches 94; Conservative 64; Mismatches 158; Indels 57; Gaps 18

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2003, 12:56:09 ; Search time 50 Seconds
(without alignments)
2761.842 Million cell updates/sec

Title: US-09-970-287-2

Perfect score: 4751

Sequence: 1 MGPPSLVLLLSATVFSILG.....PEMKRPSKSLGQLWEGWG 870

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19JUN03.*

```

1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4751	100.0	870	23	ABB80921
2	4751	100.0	885	23	ABB80922
3	4746	99.9	870	22	AAB85774
4	4746	99.9	870	22	AA179215
5	4746	99.9	870	22	AAE01440
6	4746	99.9	870	23	ABG71836
7	4746	99.9	870	23	ABG63903
8	4742	99.8	870	22	AAE01471
9	4742	99.8	870	23	ABG63902
10	4742	99.8	870	23	ABB80921
11	4719.5	99.3	867	21	AA166648
12	4719.5	99.3	867	22	AAU12181
13	4719.5	99.3	867	22	AAB65171
14	4719.5	99.3	867	23	ABB95459
15	4719.5	99.3	867	23	ABB84853
16	4719.5	99.3	867	24	ABU66579
17	4719.5	99.3	867	24	ABU66855
18	4719.5	99.3	867	24	ABU59660
19	4719.5	99.3	867	24	ABU59064
20	4719.5	99.3	867	24	ABU59211
21	4719.5	99.3	867	24	ABU59360
22	4719.5	99.3	867	24	ABU60495
23	4719.5	99.3	867	24	ABU57986
24	4719.5	99.3	867	24	ABU58917
25	4719.5	99.3	867	24	ABU13877
26	4719.5	99.3	867	24	ABU10832
27	4612	97.1	850	22	AA179216
28	4469.5	94.1	875	23	ABG71838
29	4343	91.4	818	21	ABA42268
30	3524	74.2	643	22	AA180199
31	3524	74.2	643	22	AA180200
32	3175.5	66.8	871	21	AA1800191
33	3175.5	66.8	871	22	AA185481
34	3175.5	66.8	871	23	ABG71835
35	3175.5	66.8	871	23	ABG69640
36	3175.5	66.8	871	24	ABR47506
37	3117.5	65.6	867	22	AAE00434
38	3059	64.4	818	23	AAU69417
39	3059	64.4	1611	22	AAE00438
40	2973	62.6	800	22	AA125714
41	2696	56.7	490	22	AA19319
42	2592	56.7	490	22	AA193864
43	2352	49.5	434	22	AA101522
44	2093	44.1	455	22	AA100437
45	1714.5	36.1	1114	22	ABB71505

ALIGNMENTS

```

RESULT 1
ABB80921
ID ABB80921 standard; Protein; 870 AA.
XX
AC ABB80921;
XX
DT 08-OCT-2002 (first entry)
XX
DE Human sulphatase polypeptide, 22437.
XX
KW Human; sulphatase; 22437; cytostatic; vulnerary; neuroprotective;
KW gene therapy; enzyme.
XX
OS Homo sapiens.
XX
PN WO200252019-A2.
XX
PD 04-JUL-2002.
XX
PF 03-OCT-2001; 2001WO-US30856.
XX
PR 21-DEC-2000; 2000US-257082P.
XX
(WILL-) MILLENNIUM PHARM INC.
XX
Glucksmann MA, Rudolph-Owen LA;
XX
WPI; 2002-566677/60.
XX
N-PSDB; ABB86352, ABB86353.
XX
Identifying agents for modulating (e.g. inhibiting) e.g. tumor
establishment, growth or metastases, neuron growth, or wound healing by

```

Human polypeptide
Membrane-bound pro
Human PRO1120 poly
Human PRO1120 UNQ
Human angiogenesis
Human PRO1120 prot
Human PRO polypept
Human secreted/tra
Novel secreted and
Novel human secret
Human secreted/tra
Human secreted/tra
Human PRO polypept
Human secreted/tr
Human PRO1120 poly
Human PRO polypept
Human protein SEQ
Mouse SULF2 protei
Human ORFX ORF2032
Human protein SEQ
Human protein SEQ
Breast cancer prot
Human 23553 sulfat
Human SULF1 protei
Human secreted pro
Breast cancer asso
Quail sulfatase (Q
Lung small cell ca
Human sulfatase (H
Human protein sequ
Human polypeptide,
Human polypeptide,
Human polypeptide,
Mouse gene 5 encod
Mouse sulfatase (M
Drosophila melanog

PT determining whether a test compound binds with a 22437 polypeptide
 PR (human sulfatase) -
 XX
 XX
 XX Claim 19; Fig 1A-F; 143pp; English.
 XX The invention relates to identifying a compound useful for modulating at
 CC least one phenomenon (e.g. tumour establishment, tumour growth, tumour
 CC metastases, epithelial and/or endothelial cell proliferation, neuronal
 CC cell growth, wound healing or cerebral injury). The method involves
 CC determining whether a test compound binds with a 22437 polypeptide.
 CC The identified modulators of 22437 nucleic acid and polypeptide are also
 CC useful for treating cancer or wounds (e.g. stroke-related cerebral
 CC ischaemic damage) or replacing damaged tissues. The 22437 nucleic acid
 CC and polypeptide are useful for diagnosing, preventing or treating a
 CC subject having cancer or a cellular proliferation and/or differentiation
 CC disorder or at risk of developing cancer or a cellular proliferation
 CC and/or differentiation disorder. The present sequence represents the
 CC human sulphatase polypeptide, 22437.
 XX
 SQ Sequence 870 AA;
 Query Match 100.0%; Score 4751; DB 23; Length 870;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGPPSLVCLLSATVFLSGGSAFLSHHRLKGRFDRDRNRPNIILVLTDDQVEIGS 60
 DB 1 MGPPSLVCLLSATVFLSGGSAFLSHHRLKGRFDRDRNRPNIILVLTDDQVEIGS 60
 QY 61 MQVNKTRRIMEOGGTHFINAFVTPMCCPSRSSILTKGYVHNHNTYTNENCSFSPMQA 120
 DB 61 MQVNKTRRIMEOGGTHFINAFVTPMCCPSRSSILTKGYVHNHNTYTNENCSFSPMQA 120
 QY 121 QHESRTFVNLSTGYTATFAGKVLNEYSGYPPGKWEGLLKNSRYNTLCRNGVK 180
 DB 121 QHESRTFVNLSTGYTATFAGKVLNEYSGYPPGKWEGLLKNSRYNTLCRNGVK 180
 QY 181 EKHSYSDYKDYLTDLITNDVSPFRSKNYPHRPVLWISHAAPHGSDSAPQYSRLFP 240
 DB 181 EKHSYSDYKDYLTDLITNDVSPFRSKNYPHRPVLWISHAAPHGSDSAPQYSRLFP 240
 QY 241 NASQHIPTSNYPANPKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDSMETIYNN 300
 DB 241 NASQHIPTSNYPANPKHIMRYTGMKPIHMEFTNMLQKRLQTLMSVDSMETIYNN 300
 QY 301 LVEIGELDNTYVYTABGHYHIGQFGLVKGKSNPYEFDIRVPYVRGPNVBAAGLNPHTV 360
 DB 301 LVEIGELDNTYVYTABGHYHIGQFGLVKGKSNPYEFDIRVPYVRGPNVBAAGLNPHTV 360
 QY 361 LNIDLAPTILDIAGLOIPADMGGKILKLDTERPVNRPHLKQKRVWRDSEFLVERGKLL 420
 DB 361 LNIDLAPTILDIAGLOIPADMGGKILKLDTERPVNRPHLKQKRVWRDSEFLVERGKLL 420
 QY 421 HKRNDKVDAAEENFLPKYQKVDLCORAEYOTACEOLGOKWOCVEDATGKLKHCKGP 480
 DB 421 HKRNDKVDAAEENFLPKYQKVDLCORAEYOTACEOLGOKWOCVEDATGKLKHCKGP 480
 QY 481 MRLGGSRLSNLVPKYGGSEACTCDSDGYKLSLAGRRKKLPKKYKASYVRSRSIRSV 540
 DB 481 MRLGGSRLSNLVPKYGGSEACTCDSDGYKLSLAGRRKKLPKKYKASYVRSRSIRSV 540
 QY 541 AIEVDGVRVHVLGDAAPRNLTGHWPGAPEDDDKGGDPSGTGGLPDYSAANPIKYT 600
 DB 541 AIEVDGVRVHVLGDAAPRNLTGHWPGAPEDDDKGGDPSGTGGLPDYSAANPIKYT 600
 QY 601 HRCVILENDTVQCDLDLYKSLQAWKHKLHIDHIEITLQNKIKNLREVRGHLKKRPREEC 660
 DB 601 HRCVILENDTVQCDLDLYKSLQAWKHKLHIDHIEITLQNKIKNLREVRGHLKKRPREEC 660
 QY 661 DCHKISYHTQHKRLKHGSSLPFRKGLQEKDKVLLREQKFKKRLKLLKELQNNOTC 720
 DB 661 DCHKISYHTQHKRLKHGSSLPFRKGLQEKDKVLLREQKFKKRLKLLKELQNNOTC 720

QY 721 SMEGLTCTFTHDQHWOTAPFWTLGPFCACTSANNNYWCMTINETHNLFCEPATGFE 780
 DB 721 SMEGLTCTFTHDQHWOTAPFWTLGPFCACTSANNNYWCMTINETHNLFCEPATGFE 780
 QY 781 YFDLNTDTPQLANAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840
 DB 781 YFDLNTDTPQLANAVNTLDRDVLNQLHVQLMELRSCKGYKQCNPRTRNMDLGLKDGGSYE 840
 QY 841 QYRQFORRWKPEMKRPPSSKSLGQLWEGWEG 870
 DB 841 QYRQFORRWKPEMKRPPSSKSLGQLWEGWEG 870
 RESULT 2
 ABB80922
 ID ABB80922 standard; Protein; 885 AA.
 XX
 AC ABB80922;
 XX 08-OCT-2002 (first entry)
 DT Human KIAA1427 polypeptide.
 DE Human; sulphatase; 22437; cytostatic; vulnerary; neuroprotective;
 KW gene therapy; KIAA1427.
 XX Homo sapiens.
 OS WO200252019-A2.
 PN 04-JUL-2002.
 PD 03-OCT-2001; 2001WO-US30856.
 PF 21-DEC-2000; 2000US-257082P.
 PR (MILL-) MILLENNIUM PHARM INC.
 PA Glucksmann MA, Rudolph-Owen LA;
 PI WPI; 2002-566677/60.
 DR N-PSDB; ABN86354.
 XX Identifying agents for modulating (e.g. inhibiting) e.g. tumor
 PT establishment, growth or metastases, neuron growth, or wound healing by
 PT determining whether a test compound binds with a 22437 polypeptide
 PT (human sulfatase) -
 XX Disclosure; Fig 4A-D; 143pp; English.
 PS The invention relates to identifying a compound useful for modulating at
 CC least one phenomenon (e.g. tumour establishment, tumour growth, tumour
 CC metastases, epithelial and/or endothelial cell proliferation, neuronal
 CC cell growth, wound healing or cerebral injury). The method involves
 CC determining whether a test compound binds with a 22437 polypeptide.
 CC The identified modulators of 22437 nucleic acid and polypeptide are also
 CC useful for treating cancer or wounds (e.g. stroke-related cerebral
 CC ischaemic damage) or replacing damaged tissues. The 22437 nucleic acid
 CC and polypeptide are useful for diagnosing, preventing or treating a
 CC subject having cancer or a cellular proliferation and/or differentiation
 CC disorder or at risk of developing cancer or a cellular proliferation
 CC and/or differentiation disorder. The present sequence represents a human
 CC KIAA1427 polypeptide used in alignment studies with the human sulphatase
 CC polypeptide, 22437.
 XX
 SQ Sequence 885 AA;
 Query Match 100.0%; Score 4751; DB 23; Length 885;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGPPSLVCLLSATVFLSGGSAFLSHHRLKGRFDRDRNRPNIILVLTDDQVEIGS 60
 DB 1 MGPPSLVCLLSATVFLSGGSAFLSHHRLKGRFDRDRNRPNIILVLTDDQVEIGS 60

Db 16 MGPPSLVLCILSATVFSLLGGSSAFSLSHRLKGRFQRRNRIRPNIILVLTDDQDELGS 75
 QY 61 MQVMNKRIRIMEQGGTHFINAFVTPMCCPSRSSILTGKYVHNHTYTNNECSPSQWA 120
 Db 76 MQVMNKRIRIMEQGGTHFINAFVTPMCCPSRSSILTGKYVHNHTYTNNECSPSQWA 135
 QY 121 QHESRTPAVILNSTGYRTAFPGKYLINEYNGSVPPGKWEVGLLKNRFRNYTLCRNGVK 180
 Db 136 QHESRTPAVILNSTGYRTAFPGKYLINEYNGSVPPGKWEVGLLKNRFRNYTLCRNGVK 195
 QY 191 EKHGSDYKDYLTDLITNDSVSFFRTSKMYPHRPVLVMI SHAAPHGPDSPAPQYSLRFP 240
 Db 196 EKHGSDYKDYLTDLITNDSVSFFRTSKMYPHRPVLVMI SHAAPHGPDSPAPQYSLRFP 255
 QY 241 NASOHITPSYAPNDKHWIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNM 300
 Db 256 NASOHITPSYAPNDKHWIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNM 315
 QY 301 LVETGELDNITYIVYTADHGHIQFGLVKGKSMPEYFDIRVPVYVRGPNVEAGCLNPHIV 360
 Db 316 LVETGELDNITYIVYTADHGHIQFGLVKGKSMPEYFDIRVPVYVRGPNVEAGCLNPHIV 375
 QY 361 LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVRNPHLKKQKRVWRDSFLVERGKLL 420
 Db 376 LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVRNPHLKKQKRVWRDSFLVERGKLL 435
 QY 421 HKRDNKVDQAEENFLPKYQKVKDLQCBAEYOTACEQLGQKQWQVEDATGKLKHKCKGP 480
 Db 436 HKRDNKVDQAEENFLPKYQKVKDLQCBAEYOTACEQLGQKQWQVEDATGKLKHKCKGP 495
 QY 481 MRLGGSRALSNLVPKYGGQSEACTCDSDGYKLSLAGRRKKLPKKYKASYVRSRSIRSV 540
 Db 496 MRLGGSRALSNLVPKYGGQSEACTCDSDGYKLSLAGRRKKLPKKYKASYVRSRSIRSV 555
 QY 541 AIEVDGRVYHVLGDAAPRNLTKRHWFGAPEDQDDKDGDPSTGGGLPDYSAANPIKVT 600
 Db 556 AIEVDGRVYHVLGDAAPRNLTKRHWFGAPEDQDDKDGDPSTGGGLPDYSAANPIKVT 615
 QY 601 HRCYILENDTVQCDLDLYKSLQAKNDKHLHDHEITLQNKIKNLRVGRHLKKRPEEC 660
 Db 616 HRCYILENDTVQCDLDLYKSLQAKNDKHLHDHEITLQNKIKNLRVGRHLKKRPEEC 675
 QY 661 DCHKISYHTQHKRLKRGSSLHPFRKGLQEKDKVWLLREOKRKKLKLKRLQNNDDTC 720
 Db 676 DCHKISYHTQHKRLKRGSSLHPFRKGLQEKDKVWLLREOKRKKLKLKRLQNNDDTC 735
 QY 721 SMPGLTCFTHDNQHWOTAPFTLGPFCACTSANNTYWCMTINETHNFCEPATGFLE 780
 Db 736 SMPGLTCFTHDNQHWOTAPFTLGPFCACTSANNTYWCMTINETHNFCEPATGFLE 795
 QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 840
 Db 796 YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCGYKQCNPRTRNMDLGLKDGGSYE 855
 QY 841 QYRQQRKRWPEMKRPSKSLGOLWEGWEG 870
 Db 856 QYRQQRKRWPEMKRPSKSLGOLWEGWEG 885

RESULT 3

AAB85774

ID AAB85774 standard; Protein; 870 AA.

XX

AC AAB85774;

XX

DT 29-OCT-2001 (first entry)

XX

DE Human drug metabolizing enzyme (ID No. 1558210CD1).

XX

KW Drug metabolizing enzyme; DME; immunosuppressive; cytostatic; ophthalmic;

XX

KW hepatocarcinoma; antiallergic; antiasthmatic; antibacterial; antiviral;

XX

KW antisense therapy; gene therapy; human.

XX

OS Homo sapiens.
 XX WO200159127-A2.
 XX 16-AUG-2001.
 XX 08-FEB-2001; 2001WO-US04423.
 XX 11-FEB-2000; 2000US-0181856.
 PR 17-FEB-2000; 2000US-0183684.
 PR 25-FEB-2000; 2000US-0185141.
 PR 03-MAR-2000; 2000US-0186818.
 PR 09-MAR-2000; 2000US-0188345.
 PR 17-MAR-2000; 2000US-0189997.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Tang YT, Yue H, Baughn MR, Yao MG, Bandman O, Azimzai Y, Lal P;
 PI Gandhi AR, Ring HZ, Shih LL, Yang J, Policky JU;
 XX WPI; 2001-514673/56.
 DR N-PSDB; AAH76201.
 XX Isolated polypeptide encoding a drug metabolizing enzyme useful for the
 PT diagnosis, treatment, and prevention of autoimmune/inflammatory, cell
 PT proliferative, developmental and endocrine disorders -
 XX Claim 1; Page 133-134; 150pp; English.
 CC The invention provides human drug metabolizing enzymes (DME) and
 CC polynucleotides encoding the DMEs. The DME can be expressed by standard
 CC recombinant methodology. DMEs and their agonists and antagonists are
 CC useful for the diagnosis, treatment, and prevention of autoimmune/
 CC inflammatory, cell proliferative, developmental, endocrine such as
 CC aneurysm, eye, metabolic, and gastrointestinal disorders, including liver
 CC disorders and infection. The present sequence represents a human DME.
 XX SQ Sequence 870 AA;

Query Match 99.9%; Score 4746; DB 22; Length 870;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVLCILSATVFSLLGGSSAFSLSHRLKGRFQRRNRIRPNIILVLTDDQDELGS 60
 Db 1 MGPPSLVLCILSATVFSLLGGSSAFSLSHRLKGRFQRRNRIRPNIILVLTDDQDELGS 60
 QY 61 MQVMNKRIRIMEQGGTHFINAFVTPMCCPSRSSILTGKYVHNHTYTNNECSPSQWA 120
 Db 61 MQVMNKRIRIMEQGAHFINAFVTPMCCPSRSSILTGKYVHNHTYTNNECSPSQWA 120
 QY 121 QHESRTPAVILNSTGYRTAFPGKYLINEYNGSVPPGKWEVGLLKNRFRNYTLCRNGVK 180
 Db 121 QHESRTPAVILNSTGYRTAFPGKYLINEYNGSVPPGKWEVGLLKNRFRNYTLCRNGVK 180
 QY 181 EKHGSDYKDYLTDLITNDSVSFFRTSKMYPHRPVLVMI SHAAPHGPDSPAPQYSLRFP 240
 Db 181 EKHGSDYKDYLTDLITNDSVSFFRTSKMYPHRPVLVMI SHAAPHGPDSPAPQYSLRFP 240
 QY 241 NASOHITPSYAPNDKHWIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNM 300
 Db 241 NASOHITPSYAPNDKHWIMRYTGPMPKPIHMEFTNMLQKRLQTLMSVDDSMETIYNM 300
 QY 301 LVETGELDNITYIVYTADHGHIQFGLVKGKSMPEYFDIRVPVYVRGPNVEAGCLNPHIV 360
 Db 301 LVETGELDNITYIVYTADHGHIQFGLVKGKSMPEYFDIRVPVYVRGPNVEAGCLNPHIV 360
 QY 361 LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVRNPHLKKQKRVWRDSFLVERGKLL 420
 Db 361 LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVRNPHLKKQKRVWRDSFLVERGKLL 420
 QY 421 HKRDNKVDQAEENFLPKYQKVKDLQCBAEYOTACEQLGQKQWQVEDATGKLKHKCKGP 480

Db 421 HKRDNDKVDQAEEENFLPKYQRYKDLQRAEYQTACEQLGQKWQCVEDATGKLKHKCKGP 480

QY 481 MRLGGSRALSNLVPKYGGSGSEACTCDSDYKLSLAGREKLLFKKKYKASYRSRSIRSV 540

Db 481 MRLGGSRALSNLVPKYGGSGSEACTCDSDYKLSLAGREKLLFKKKYKASYRSRSIRSV 540

QY 541 AIEVDGRVYHVLGDAQAOPRNLTKRHWPAGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600

Db 541 AIEVDGRVYHVLGDAQAOPRNLTKRHWPAGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600

QY 601 HRCYILENDTVQCDLDLYKSLQAWKHKLHIDHETLQNKIKNLRVRGHLKKRPEEC 660

Db 601 HRCYILENDTVQCDLDLYKSLQAWKHKLHIDHETLQNKIKNLRVRGHLKKRPEEC 660

QY 661 DCHKISYHTQHKGRGLKRGSSLHPPFRKGLQKDKVWLLREQKXKKLKLKRLQNNDTG 720

Db 661 DCHKISYHTQHKGRGLKRGSSLHPPFRKGLQKDKVWLLREQKXKKLKLKRLQNNDTG 720

QY 721 SMPGLTCTFTHDNOHWOTAPFWTLGPPCACTSANNTYWCMTINETHNLFCEFATGFE 780

Db 721 SMPGLTCTFTHDNOHWOTAPFWTLGPPCACTSANNTYWCMTINETHNLFCEFATGFE 780

QY 781 YEDLNTDYPQLMNAVNTLDROVLNQLHVQLMELRSCKGYKQCNPRTRMDLGLKGGSYE 840

Db 781 YEDLNTDYPQLMNAVNTLDROVLNQLHVQLMELRSCKGYKQCNPRTRMDLGLKGGSYE 840

QY 841 QYRQFQRKWPMPKRPSSKSLGQWEGWEG 870

Db 841 QYRQFQRKWPMPKRPSSKSLGQWEGWEG 870

RESULT 4

ID AAM79215 standard; Protein; 870 AA.

XX AC AAM79215;

XX DT 06-NOV-2001 (first entry)

XX DE Human protein SEQ ID NO 1877.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;

XX KW nervous system disorder; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US04098.

XX PR 03-FEB-2000; 2000US-0496914.

XX PR 27-APR-2000; 2000US-0560875.

XX PR 20-JUN-2000; 2000US-0598075.

XX PR 19-JUL-2000; 2000US-0620325.

XX PR 01-SEP-2000; 2000US-0654936.

XX PR 15-SEP-2000; 2000US-0663561.

XX PR 20-OCT-2000; 2000US-0693325.

XX PR 30-NOV-2000; 2000US-0728422.

XX PA (HYSE-) HYSEQ INC.

XX PI Zhao YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI; 2001-476283/51.

XX DR N-PSDB; AAK52348.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities,

PT useful in diagnosis and gene therapy -

XX Claim 20; Page 4262-4264; 622lpp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication.

XX SQ Sequence 870 AA;

Query Match 99.9%; Score 4746; DB 22; Length 870;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVLCLLSATVFSLLGGSSAFLSHRLKGRFORDRNRIRPNILVLTDDQDVELGS 60

Db 1 MGPPSLVLCLLSATVFSLLGGSSAFLSHRLKGRFORDRNRIRPNILVLTDDQDVELGS 60

QY 61 MQVMNKTTRIMEOGGTHFINAFVTTMCCPSRSSILTKYVHNHNTYTNENCSFSSQWA 120

Db 61 MQVMNKTTRIMEOGGAHFINAFVTTMCCPSRSSILTKYVHNHNTYTNENCSFSSQWA 120

QY 121 QHESRTFAVYLNSTGYRTAFFGKYLYNEINGSYPPGKWGWGLLKNSRFNYTLCRGVK 180

Db 121 QHESRTFAVYLNSTGYRTAFFGKYLYNEINGSYPPGKWGWGLLKNSRFNYTLCRGVK 180

QY 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGSDAPQYSLRFP 240

Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGSDAPQYSLRFP 240

QY 241 NASOHITPSYNYAPNPKHMYRTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300

Db 241 NASOHITPSYNYAPNPKHMYRTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300

QY 301 LVETGELDNTYIVYTADHGVHIGQGLVKGKSMPEYFDIRVPEYVGGNVGACLNPHIV 360

Db 301 LVETGELDNTYIVYTADHGVHIGQGLVKGKSMPEYFDIRVPEYVGGNVGACLNPHIV 360

QY 361 LNIDLAPTILDIAGLDIPADMKGKSLKLLDTERPVNRFHLKKMVRWDRSFLVERGKLL 420

Db 361 LNIDLAPTILDIAGLDIPADMKGKSLKLLDTERPVNRFHLKKMVRWDRSFLVERGKLL 420

QY 421 HKRDNDKVDQAEEENFLPKYQRYKDLQRAEYQTACEQLGQKWQCVEDATGKLKHKCKGP 480

Db 421 HKRDNDKVDQAEEENFLPKYQRYKDLQRAEYQTACEQLGQKWQCVEDATGKLKHKCKGP 480

QY 481 MRLGGSRALSNLVPKYGGSGSEACTCDSDYKLSLAGREKLLFKKKYKASYRSRSIRSV 540

Db 481 MRLGGSRALSNLVPKYGGSGSEACTCDSDYKLSLAGREKLLFKKKYKASYRSRSIRSV 540

QY 541 AIEVDGRVYHVLGDAQAOPRNLTKRHWPAGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600

Db 541 AIEVDGRVYHVLGDAQAOPRNLTKRHWPAGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600

QY 601 HRCYILENDTVQCDLDLYKSLQAWKHKLHIDHETLQNKIKNLRVRGHLKKRPEEC 660

Db 601 HRCYILENDTVQCDLDLYKSLQAWKHKLHIDHETLQNKIKNLRVRGHLKKRPEEC 660

QY 661 DCHKISYHTQHKGRGLKRGSSLHPPFRKGLQKDKVWLLREQKXKKLKLKRLQNNDTG 720

Db 661 DCHKISYHTQHKGRGLKRGSSLHPPFRKGLQKDKVWLLREQKXKKLKLKRLQNNDTG 720

QY 721 SMPGLTCTFDHONQWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLECFEPATGFL 780
 Db 721 SMPGLTCTFDHONQWOTAPFWTLGPFCACTSANNTYWCMTINETHNFLECFEPATGFL 780
 QY 781 YFDLNTDPYQLMNAVNTLDRVLNQLHVQLMELASCKGYKQCNPRTRNMGLGKDGGSYE 840
 Db 781 YFDLNTDPYQLMNAVNTLDRVLNQLHVQLMELASCKGYKQCNPRTRNMGLGKDGGSYE 840
 QY 841 QYRQPRKWPCKPSSKSLGQLWEGWEG 870
 Db 841 QYRQPRKWPCKPSSKSLGQLWEGWEG 870
 RESULT 5
 AAE01440
 ID AAE01440 standard; Protein; 870 AA.
 XX AC
 XX AC
 XX DT 17-JUL-2001 (first entry)
 XX DE Human gene 5 encoded secreted protein HE9QN39, SEQ ID NO:95.
 XX KW Human; secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnerability;
 KW cell culture; chemotaxis; food additive; gene therapy;
 KW binding partner identification.
 XX OS Homo sapiens.
 XX FH Key
 XX FT Peptide
 XX FT 1..24
 XX FT /label= Signal_peptide
 XX FT Protein
 XX FT 25..870
 XX FT /note= "Mature human secreted protein"
 XX PN WO200134626-A1.
 XX PD 17-MAY-2001.
 XX PF 01-NOV-2000; 2000WO-US30045.
 XX PF 05-NOV-1999; 98US-0163581.
 XX PR 30-JUN-2000; 2000US-0215133.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Ruben SM, Komatsoulis GA, Moore PA, Birse CE, Ni J;
 XX WPI; 2001-308778/32.
 XX DR N-PSDB; AAD05304.
 XX PT New nucleic acid molecules encoding 28 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives -
 XX PS Claim 11; Page 486-489; 562pp; English.
 XX AAD05300-AAD05379 represent cDNAs corresponding to 28 human secreted
 CC protein genes, and AAE01436-AAE01513 represent the proteins or variants.
 CC AAE01514-AAE01544 represent human secreted protein fragments or variants.
 CC The genes and their secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the

CC 28 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein of the invention.
 XX SQ Sequence 870 AA;

Query Match 99.9%; Score 4746; DB 22; Length 870;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGPPSLVLCLLSATVFSLLGGSSAFSLHRLKGRFQDRNRIRPNIILVLTDDQDVELGS 60
 Db 1 MGPPSLVLCLLSATVFSLLGGSSAFSLHRLKGRFQDRNRIRPNIILVLTDDQDVELGS 60
 QY 61 MQVMKTRIRIMEOGTHFINAFVTPMCCPSRSILLTKYVHNHTYNNENCSPPSHQA 120
 Db 61 MQVMKTRIRIMEOGGAHFINAFVTPMCCPSRSILLTKYVHNHTYNNENCSPPSHQA 120
 QY 121 QHESRTFAVYLNSTGYRTAFFGKYLYNEVNGSVVPPGKWEVGLLKNRSFYNTLCRNVK 180
 Db 121 QHESRTFAVYLNSTGYRTAFFGKYLYNEVNGSVVPPGKWEVGLLKNRSFYNTLCRNVK 180
 QY 181 EKHSYDYSKDYLTDLITNDSVSFFRTSKMTYPPHRLMVI SHAAHPGDESAPOYSRLFP 240
 Db 181 EKHSYDYSKDYLTDLITNDSVSFFRTSKMTYPPHRLMVI SHAAHPGDESAPOYSRLFP 240
 QY 241 NASQHTPSYVYAPNDKHWIMRYTGPMKPIMHEFTNMLQKRLQTLMSVDDSMETIYNN 300
 Db 241 NASQHTPSYNTAPNDKHWIMRYTGPMKPIMHEFTNMLQKRLQTLMSVDDSMETIYNN 300
 QY 301 LVETGELDNTYIVYTADHGYHIGQGLVKGXSMPEYFDIRVPFVYRGNVEAGCLNPHIV 360
 Db 301 LVETGELDNTYIVYTADHGYHIGQGLVKGXSMPEYFDIRVPFVYRGNVEAGCLNPHIV 360
 QY 361 LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVNRFHLKQKQVWRDSFLVERGKLL 420
 Db 361 LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVNRFHLKQKQVWRDSFLVERGKLL 420
 QY 421 HKRDNKDYDAOBENFLPKYQVYKDLQRAEYQTAECQGLQKQWQVEDATGKLKHKCKGP 480
 Db 421 HKRDNKDYDAOBENFLPKYQVYKDLQRAEYQTAECQGLQKQWQVEDATGKLKHKCKGP 480
 QY 481 MRLGGSRALSNLVPKYQOGSEACTCDSDGYKLSLAGRRKLFKKYKASYRSRSIRSV 540
 Db 481 MRLGGSRALSNLVPKYQOGSEACTCDSDGYKLSLAGRRKLFKKYKASYRSRSIRSV 540
 QY 541 ALEVDGRVYHVLGDAAQPRNITKRWHPGAPEDODDKGDFSGTGLPDYSAANPKVT 600
 Db 541 ALEVDGRVYHVLGDAAQPRNITKRWHPGAPEDODDKGDFSGTGLPDYSAANPKVT 600
 QY 601 HRCYILENDTVQCDDLDLYKSLQAWKHKLHIDHETLQNKIKLREVRGHLKKRPEEC 660
 Db 601 HRCYILENDTVQCDDLDLYKSLQAWKHKLHIDHETLQNKIKLREVRGHLKKRPEEC 660
 QY 661 DCHKISYTHQHKRLKHRGSSLHPFRKGLQEKQVWLLREQKRXKRLKLLKRLQNNDTC 720

Db 661 DCHKISYHTQHKGRKLRSSLPFRKGLQEKDQWVLRQKQKQKRLKRLQNNQTC 720
QY 721 SMPGLTCTFHDNQHWQAPFTWLTGPPFCACTSANNNTYWCMRTINETHNPLFCEPATGFLE 780
Db 721 SMPGLTCTFHDNQHWQAPFTWLTGPPFCACTSANNNTYWCMRTINETHNPLFCEPATGFLE 780
QY 781 YFDLNTDPPQLMNAVNTLDRVNLQHLVQLMELRSCGKYKQCNPRTRMDLGLKDGGSYE 840
Db 781 YFDLNTDPPQLMNAVNTLDRVNLQHLVQLMELRSCGKYKQCNPRTRMDLGLKDGGSYE 840
QY 841 QYRQFORRWPEWKPSSKSLQGLWEGWEG 870
Db 841 QYRQFORRWPEWKPSSKSLQGLWEGWEG 870

RESULT 6
ABG71836
ID ABG71836 standard; Protein; 870 AA.

XX AC ABG71836;
XX DT 20-JAN-2003 (first entry)
XX DE human SULF2 protein.

XX KW Human; SULF2; glucosamine-6-sulphatase; cancer; ischaemia; enzyme;
KW tumour; angiogenesis; coronary; carotid; arterial occlusive disease;
KW peripheral arterial disease; atherosclerosis; myointimal hyperplasia;
KW thromboangitis obliterans; thrombotic disorder; vasculitis;
KW heart attack; myocardial infarction; vascular death; inflammation;
KW rheumatoid arthritis; asthma; adult respiratory distress syndrome;
KW sarcoidosis; hypersensitivity pneumonitis; multiple sclerosis;
KW allograft rejection; lymphoma; thrombosis; sulphatase.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
FT Peptide 1..22
FT Protein /note= "Signal peptide"
FT Cleavage-site /note= "Mature protein"
FT Cleavage-site /note= "furan/PACE protease cleavage site"
FT Cleavage-site /note= "furan/PACE protease cleavage site"
FT Cleavage-site /note= "furan/PACE protease cleavage site"
FT Cleavage-site /note= "furan/PACE protease cleavage site"

XX WO200259327-A2.

XX PD 01-AUG-2002.

XX PF 26-DEC-2001; 2001WO-US49793.

XX PR 27-DEC-2000; 2000US-258577P.

XX PR 09-FEB-2001; 2001US-267831P.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Rosen SD, Hemmerich S, Tomita M, Palmeri D;

XX DR WPI: 2002-636541/68.

XX DR N-PSDB; ABS56553.

XX PT New sulfate polypeptides having glucosamine-6-sulfatase activity,
PT useful in screening, discovery and prepatation of diagnostic and
PT therapeutic agents for treating cancer, ischemic conditions,
PT inflammation, or thrombosis

XX PS Example 1; Fig2A; 293pp; English.

XX CC This invention relates to the DNA and protein sequences of a novel
CC polypeptide having glucosamine-6-sulphatase activity. The sulphatases of
CC are useful in screening, discovery and prepatation of diagnostic and
CC therapeutic agents for treating cancer, ischaemic conditions,
CC inflammation, or thrombosis. The nucleic acids are useful in preparing
CC the sulfate polypeptides, identifying the expression of genes in a
CC biological specimen, or generating transgenic non-human animals or
CC site-specific gene modification in cell lines. The host cells are
CC useful in replicating and/or expressing the polynucleotides or nucleic
CC acids. The agents are useful in treating the disorders cited above by
CC reducing tumour growth, inflammation, and thrombosis, or increasing
CC angiogenesis, e.g. by treating coronary, carotid, or arterial occlusive
CC disease, peripheral arterial disease, atherosclerosis, myointimal
CC hyperplasia, thromboangitis obliterans, thrombotic disorders,
CC vasculitis; or preventing ischemic conditions, heart attack (myocardial
CC infarction), or other vascular death. The sulphatases and/or agents are
CC also useful in treating rheumatoid arthritis, asthma, adult respiratory
CC distress syndrome, sarcoidosis, hypersensitivity pneumonitis, multiple
CC sclerosis, allograft rejection, and spread of lymphomas to cutaneous
CC sites. The present sequence represents the human SULF2 protein of the
CC invention.
XX CC
SQ Sequence 870 AA;

Query Match 99.9%; Score 4746; DB 23; Length 870;

Best Local Similarity 99.9%; Pred. NO. 0;

Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFORDRNRNPNIILVLTDDQDELGS 60

Db 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFORDRNRNPNIILVLTDDQDELGS 60

Qy 61 MQVNNKTRIRMEQGGTHPINAFTTMCPCPSRSSILTKYVHNHNTYTNENCSSPSWA 120

Db 61 MQVNNKTRIRMEQGGTHPINAFTTMCPCPSRSSILTKYVHNHNTYTNENCSSPSWA 120

Qy 121 QHESRTPAVYLNSTGYRTAFFGKYLNEYNGSYVPPGKWEVGLLKNSRFYNYTLCRNGYK 180

Db 121 QHESRTPAVYLNSTGYRTAFFGKYLNEYNGSYVPPGKWEVGLLKNSRFYNYTLCRNGYK 180

Qy 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPLVWLVISHAAPHGSDAPQYSLRFP 240

Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPLVWLVISHAAPHGSDAPQYSLRFP 240

Qy 241 NASOHLTPSYNYPNDPKWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300

Db 241 NASOHLTPSYNYPNDPKWIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300

Qy 301 LVETGELDNTYIYVTADHGYHIGQFGLVKGSMPYEFDIRVPYVVRGPNVVEAGCLNPHIV 360

Db 301 LVETGELDNTYIYVTADHGYHIGQFGLVKGSMPYEFDIRVPYVVRGPNVVEAGCLNPHIV 360

Qy 361 LNIDLAPTILDIAGLOIPADMCKSLKLLDTERPVNRPHLKQOMRVWRDSEFLVERGKLL 420

Db 361 LNIDLAPTILDIAGLOIPADMCKSLKLLDTERPVNRPHLKQOMRVWRDSEFLVERGKLL 420

Qy 421 HKRDNKVDQAQENFLPKYQORVKDLCORAEYQACQLGQKQCVEDATCKLKHCKGP 480

Db 421 HKRDNKVDQAQENFLPKYQORVKDLCORAEYQACQLGQKQCVEDATCKLKHCKGP 480

Qy 481 MRLGGSRALSNLVPKYGGGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540

Db 481 MRLGGSRALSNLVPKYGGGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540

Qy 541 AIEVDGRVHYVGLGDAAPRNLTKRHWPGAPEDDDKDGDFSGTGGLPDYSAANPIKVT 600

Db 541 AIEVDGRVHYVGLGDAAPRNLTKRHWPGAPEDDDKDGDFSGTGGLPDYSAANPIKVT 600

Qy 601 HRCVILENDTVQCDLDLYKSLQAKDKHLHDHIEIETLQNKILREVGRHLKCKGPEEC 660

Db 601 HRCVILENDTVQCDLDLYKSLQAKDKHLHDHIEIETLQNKILREVGRHLKCKGPEEC 660

QY 661 DCHKISYHTQHKRLKXRGSSLPFRKGLQEKDQVWLLRQKRRKKLRLKRLQNNDT 720
 D 661 DCHKISYHTQHKRLKXRGSSLPFRKGLQEKDQVWLLRQKRRKKLRLKRLQNNDT 720
 QY 721 SMPGLTCFTHDNQHMOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCERATG 780
 D 721 SMPGLTCFTHDNQHMOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCERATG 780
 QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVLQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
 D 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVLQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
 QY 841 QYQFORRWKPEMKRPPSSKSLQOLWEGWEG 870
 D 841 QYQFORRWKPEMKRPPSSKSLQOLWEGWEG 870

RESULT 7
 ABG63903

ID ABG63903 standard; Protein; 870 AA.

AC ABG63903;

XX 27-AUG-2002 (first entry)

DT Human albumin fusion protein #578.

XX Albumin fusion protein; therapeutic protein X; human albumin; HA;
 KW human serum albumin; HSA; cancer; reproductive disorder;
 KW digestive disorder; immune disorder; endocrine disorder;
 KW haematopoietic disorder; neural disorder; connective disorder;
 KW cytotatic; antiinfertility; antiinflammatory; antiulcer;
 KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
 KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
 KW osteopathic; antiarthritic.

XX Homo sapiens.

OS Synthetic.

XX WO20017137-A1.

PN 18-OCT-2001.

XX 12-APR-2001; 2001WO-US11988.

XX 12-APR-2000; 2000US-229358P.

XX 25-APR-2000; 2000US-199384P.

XX 21-DEC-2000; 2000US-256931P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Haseltine WA;

PI WPI; 2002-010886/01.

XX New fusion protein for treating disease e.g. diabetes comprises an
 PT albumin fused to a therapeutic protein -
 XX Claim 1; Page 892-894; 2102pp; English.

PS The present invention relates to albumin fusion proteins comprising a
 CC therapeutic protein X and human albumin (HA, also known as human serum
 CC albumin, HSA). The proteins are useful for treating a disease or
 CC disorder that may be modulated by therapeutic protein X. The albumin
 CC extends the shelf-life of protein X, and may increase its biological
 CC in vitro/in vivo activity. The protein is useful for treating and
 CC diagnosing disorders such as cancer, reproductive disorders, digestive
 CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
 CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
 CC (e.g. diabetes), haematopoietic disorders, neural disorders
 CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
 CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
 CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin

CC fusion proteins of the invention.
 XX
 SQ Sequence 870 AA;

Query March 99.9%; Score 4746; DB 23, Length 870;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFORRRNIRPNILVLTDODDELGS 60
 D 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFORRRNIRPNILVLTDODDELGS 60
 QY 61 MQVMNTRIMEQGGTHFINAFVTTMCCPSRSSILTOKYVHNNTYNNENCSPPSQA 120
 D 61 MQVMNTRIMEQGGAHFINAFVTTMCCPSRSSILTOKYVHNNTYNNENCSPPSQA 120
 QY 121 QHESRTFAVLNSTGYRTAFGKYLNEVNGSVYPPGKWKVGLLKNRSFYNTLCRNGVK 180
 D 121 QHESRTFAVLNSTGYRTAFGKYLNEVNGSVYPPGKWKVGLLKNRSFYNTLCRNGVK 180
 QY 181 EXHGSYSDYLTDLITNDSVSFFRTSKMYPHRLVMI SHAAPHGPEDSAPQYSRLFP 240
 D 181 EXHGSYSDYLTDLITNDSVSFFRTSKMYPHRLVMI SHAAPHGPEDSAPQYSRLFP 240
 QY 241 NASOHLTPSYNAPNDKHMIMRYTGMKPIHMEFTNMLQKRRLQTLMSVDDSMETIYNN 300
 D 241 NASOHLTPSYNAPNDKHMIMRYTGMKPIHMEFTNMLQKRRLQTLMSVDDSMETIYNN 300
 QY 301 LVETGELONTYIVYADHGYHIGFGLVKGKSMPEYEDIRVPFYVYRGPNVEAGCLNPHIV 360
 D 301 LVETGELONTYIVYADHGYHIGFGLVKGKSMPEYEDIRVPFYVYRGPNVEAGCLNPHIV 360
 QY 361 LNIIDLAPTILDTAGLDIPADMDGKSLILKLDTERPVNRFHKKKRWVRDSEFLVERGKLL 420
 D 361 LNIIDLAPTILDTAGLDIPADMDGKSLILKLDTERPVNRFHKKKRWVRDSEFLVERGKLL 420
 QY 421 HKRDNDKVDQAQENFLPKYQYRVCOLCORAEYQVTCGLQKQKQWOCVEDATGKLHKCKGP 480
 D 421 HKRDNDKVDQAQENFLPKYQYRVCOLCORAEYQVTCGLQKQKQWOCVEDATGKLHKCKGP 480
 QY 481 MRLGGSRALSNLVPKYGGGSEACTCDSDGYKLSLAGRRKKLFKKYKASYVRSRSRSV 540
 D 481 MRLGGSRALSNLVPKYGGGSEACTCDSDGYKLSLAGRRKKLFKKYKASYVRSRSRSV 540
 QY 541 AIEVDGRVYHVGLDAAOPRNLTKRHMPGAPEDDDKDGDFSGTGGLPDYSAANPKVT 600
 D 541 AIEVDGRVYHVGLDAAOPRNLTKRHMPGAPEDDDKDGDFSGTGGLPDYSAANPKVT 600
 QY 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHDHJETLQNKIKNLREVRGHLKKRPREEC 660
 D 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHDHJETLQNKIKNLREVRGHLKKRPREEC 660
 QY 661 DCHKISYHTQHKRLKXRGSSLPFRKGLQEKDQVWLLRQKRRKKLRLKRLQNNDT 720
 D 661 DCHKISYHTQHKRLKXRGSSLPFRKGLQEKDQVWLLRQKRRKKLRLKRLQNNDT 720
 QY 721 SMPGLTCFTHDNQHMOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCERATG 780
 D 721 SMPGLTCFTHDNQHMOTAPFWTLGPFCACTSANNTYWCMTINETHNPLFCERATG 780
 QY 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVLQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
 D 781 YFDLNTDPYQLMNAVNTLDRDVLNQLHVLQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE 840
 QY 841 QYQFORRWKPEMKRPPSSKSLQOLWEGWEG 870
 D 841 QYQFORRWKPEMKRPPSSKSLQOLWEGWEG 870

RESULT 8
 AAE01471
 ID AAE01471 standard; Protein; 870 AA.
 XX

AC AAE01471;
 DT 17-JUL-2001 (first entry)
 XX Human gene 5 encoded secreted protein HE9QN39, SEQ ID NO:127.
 XX Human; secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiodysplasia; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnery;
 KW cell culture; chemotaxis; food additive; gene therapy;
 KW binding partner identification.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..24
 FT /label= Signal_peptide
 FT Protein 25..870
 FT /note= "Mature human secreted protein"
 FT Misc-difference 537
 FT /label= Unknown
 FT /note= "Encoded by ART"
 XX WO200134626-A1.
 PN
 PD 17-MAY-2001.
 PD
 PD 01-NOV-2000; 2000WO-US30045.
 PF
 XX 05-NOV-1999; 99US-0163581.
 PR 30-JUN-2000; 2000US-0215133.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Komatsoulis GA, Moore PA, Birse CE, Ni J;
 PI WP: 2001-308778/32.
 DR N-PSDB; AAD01471.
 DR
 XX New nucleic acid molecules encoding 28 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives -
 XX
 PS Claim 11; Page 507-510; 562pp; English.
 CC AAD05300-AAD05379 represent cDNAs corresponding to 28 human secreted
 CC protein genes, and AAE01436-AAE01513 represent the proteins they encode.
 CC AAE01514-AAE01544 represent human secreted protein fragments or variants.
 CC The genes and their secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 28 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiodysplasia, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their

CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein of the invention.
 XX
 SQ Sequence 870 AA;
 Query Match 99.8%; Score 4742; DB 22; Length 870;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 868; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 MGPPSLVCLLSATVPSLLGGSSAPLSHRLKGRFDRNRINPNILVLTDQDVELGS 60
 Db 1 MGPPSLVCLLSATVPSLLGGSSAPLSHRLKGRFDRNRINPNILVLTDQDVELGS 60
 Qy 61 MQVMNKTTRIMEGGTHFINAFVTTMCCPSRSSILTKYVHNHTYTNENCSSPSWQA 120
 Db 61 MQVMNKTTRIMEGGGAHFINAFVTTMCCPSRSSILTKYVHNHTYTNENCSSPSWQA 120
 Qy 121 QHESRTFAVYLNSTGYRTAFFGKYLYNEVNGSYVPPGKEMWGLLKNSRFNYTLCENGVK 180
 Db 121 QHESRTFAVYLNSTGYRTAFFGKYLYNEVNGSYVPPGKEMWGLLKNSRFNYTLCENGVK 180
 Qy 181 EKIGSDYSKDYLTDLITNDSVSFFRTSKOMYHPRPVLMVISHAAPHGSDSAPQSRLEP 240
 Db 181 EKIGSDYSKDYLTDLITNDSVSFFRTSKOMYHPRPVLMVISHAAPHGSDSAPQSRLEP 240
 Qy 241 NASQHITPSYNYAPNPKWIMRYTGPMPKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300
 Db 241 NASQHITPSYNYAPNPKWIMRYTGPMPKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300
 Qy 301 LVETGELDNTYIVYTADHGVIHQFGLVKGKSMPEYFDIRVPPYVRGPNVEAGLNPHIV 360
 Db 301 LVETGELDNTYIVYTADHGVIHQFGLVKGKSMPEYFDIRVPPYVRGPNVEAGLNPHIV 360
 Qy 361 LNTDLAPTLDIAGLDIPADMDGKSLKLLDTERPVRNPHLKKGRVWRDSEFLVERGKLL 420
 Db 361 LNTDLAPTLDIAGLDIPADMDGKSLKLLDTERPVRNPHLKKGRVWRDSEFLVERGKLL 420
 Qy 421 HKEDNDKVDAAQENFLPKYQVRVDLCQRAYOTACEOLGQWQCVEDATGKLKHKCKGP 480
 Db 421 HKEDNDKVDAAQENFLPKYQVRVDLCQRAYOTACEOLGQWQCVEDATGKLKHKCKGP 480
 Qy 481 MRLGGSRALSNLVPKYVGQSSACTCDSDGYKLSLAGRRKLPKKYKASYVRSRSIRSV 540
 Db 481 MRLGGSRALSNLVPKYVGQSSACTCDSDGYKLSLAGRRKLPKKYKASYVRSRSIRSV 540
 Qy 541 AIEVDGRVHVGLGDAAPRNLTTRHWPAGAPEDDDKGGDFSGTGGLPDYSAANPIKYT 600
 Db 541 AIEVDGRVHVGLGDAAPRNLTTRHWPAGAPEDDDKGGDFSGTGGLPDYSAANPIKYT 600
 Qy 601 HRCYILENDTVQCDLKYSLQAKDHLHIDHEIETLQNKIKNREVGRHLKKRPEEC 660
 Db 601 HRCYILENDTVQCDLKYSLQAKDHLHIDHEIETLQNKIKNREVGRHLKKRPEEC 660
 Qy 661 DCHKISYTHQKRLKHGSSLLHPFKGLQKQKDWLLREQKQKLLKRLKRLQNNDTFC 720
 Db 661 DCHKISYTHQKRLKHGSSLLHPFKGLQKQKDWLLREQKQKLLKRLKRLQNNDTFC 720
 Qy 721 SMPGLTCFTHDNHQTAPFWTLGPPCACTSANNTYCMRTINETHNLFCEFAFGFUE 780
 Db 721 SMPGLTCFTHDNHQTAPFWTLGPPCACTSANNTYCMRTINETHNLFCEFAFGFUE 780
 Qy 781 YFDLNTDPQLMANVNTLDRDLNQLHVLQMLRSCKYKQCNPRTRNMDLGLKGGSYE 840
 Db 781 YFDLNTDPQLMANVNTLDRDLNQLHVLQMLRSCKYKQCNPRTRNMDLGLKGGSYE 840
 Qy 841 QYRQFORRWKPEMKRPSKSLQQLMEGWEG 870
 Db 841 QYRQFORRWKPEMKRPSKSLQQLMEGWEG 870

QY	121	QHEGRTPAVYLNSTGYETAFFGKYLNEYNGSYVPPGKQKQWVGLLKNRSRPNTYLCRNGVK	180
Db	121	QHEGRTPAVYLNSTGYETAFFGKYLNEYNGSYVPPGKQKQWVGLLKNRSRPNTYLCRNGVK	180
QY	181	EKHGSDYSKDYLTDLITNDVSVPFRSTSKMYPHRPVLWVISHAAPHGPDSDAPQYSRLFP	240
Db	181	EKHGSDYSKDYLTDLITNDVSVPFRSTSKMYPHRPVLWVISHAAPHGPDSDAPQYSRLFP	240
QY	241	NASQHIPTSYNYAENPKGHIIMRYTGPMKPIHMEFTNMLQRKLOTLMSVDDSMETIYNN	300
Db	241	NASQHIPTSYNYAENPKGHIIMRYTGPMKPIHMEFTNMLQRKLOTLMSVDDSMETIYNN	300
QY	301	LVEGTGELDNTYIVVTADHGVIHQFGLVKGSKMPYEFDIRVPPYVRGPNVAGCLNPHIV	360
Db	301	LVEGTGELDNTYIVVTADHGVIHQFGLVKGSKMPYEFDIRVPPYVRGPNVAGCLNPHIV	360
QY	361	LNIDLAFTIIDIAGLDIPADMDGKSILKLLDTERPVNRPHLKQKQVWRDSDFLVERGKLL	420
Db	361	LNIDLAFTIIDIAGLDIPADMDGKSILKLLDTERPVNRPHLKQKQVWRDSDFLVERGKLL	420
QY	421	HKRDNNDKVDAAQEEENFLPKYQRVXDLQORAEYQTAQEQLGQKQWQVEDATGKLLHKCKGP	480
Db	421	HKRDNNDKVDAAQEEENFLPKYQRVXDLQORAEYQTAQEQLGQKQWQVEDATGKLLHKCKGP	480
QY	481	MRLGGSRALSNLVPKYYGGSEACTCDSGDYKLSLAGRRKKLPKKYKASYVRSRSIRSV	540
Db	481	MRLGGSRALSNLVPKYYGGSEACTCDSGDYKLSLAGRRKKLPKKYKASYVRSRSIRSV	540
QY	541	ALEVDGVRVYHVLGDAQAQPNRLTKRHHWPGAPEDDKDGDFSGTGGLPDYSAANPIKVT	600
Db	541	ALEVDGVRVYHVLGDAQAQPNRLTKRHHWPGAPEDDKDGDFSGTGGLPDYSAANPIKVT	600
QY	601	HRCYILENDTVQCDLDLYKSLQAWKHKLHIHIEITLQNKIKNLREVRGHLKKKQPEEC	660
Db	601	HRCYILENDTVQCDLDLYKSLQAWKHKLHIHIEITLQNKIKNLREVRGHLKKKQPEEC	660
QY	661	DCHKISYHTQHKGRKLRGSSLPFRKGLQEKQKVMLLRQKHKKLKLLKRLQNNDTG	720
Db	661	DCHKISYHTQHKGRKLRGSSLPFRKGLQEKQKVMLLRQKHKKLKLLKRLQNNDTG	720
QY	721	SNPGLTCFTHDNQHWQTAPEWTLGGPFCACTSANNNTYWCMTINETHNFIJCFEFATGPLE	780
Db	721	SNPGLTCFTHDNQHWQTAPEWTLGGPFCACTSANNNTYWCMTINETHNFIJCFEFATGPLE	780
QY	781	YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTRNMDLGLKDGGSYE	840
Db	781	YFDLNTDPYQLMNAVNTLDRDVLNQLHVQLMELRSCKYKQCNPRTRNMDLGLKDGGSYE	840
QY	841	QYRQFQRKRWPEMKRPPSKSLGQLEWEGW 870	
Db	841	QYRQFQRKRWPEMKRPPSKSLGQLEWEGW 870	
RESULT 10			
ABB90378			
XX	ID	ABB90378 standard; Protein; 870 AA.	
XX	AC	ABB90378;	
XX	DT	24-MAY-2002 (first entry)	
XX	DE	Human polypeptide SEQ ID NO 2754.	
XX	DE		
KW	KW	Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;	
KW	KW	antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;	
KW	KW	vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;	
KW	KW	cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;	
KW	KW	neurological disease; infection; human; secreted protein.	
XX	OS	Homo sapiens.	
XX	PN	WO200190304-A2.	

29-NOV-2001.
18-MAY-2001; 2001WO-US16450.
19-MAY-2000; 2000US-205515P.
(HUMA-) HUMAN GENOME SCI INC.
Birse CE, Rosen CA;
WPI: 2002-122018/16.
N-PSDB; ABL90787.
Novel 1405 isolated polypeptides, useful for diagnosis, treatment and prevention of neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders -
Claim 11; SEQ ID NO 2754; 2081pp + Sequence Listing; English.
The invention relates to novel genes (ABL89449-ABL90853) and proteins (ABB89040-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Query Match 99.8%; Score 4742; DB 23; Length 870;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 868; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGPPSLVLCCLLSATVPSLLGGSSAFUSHRLKGRFORRRNIRPNILVLTDQDVELGS 60
DB 1 MGPPSLVLCCLLSATVPSLLGGSSAFUSHRLKGRFORRRNIRPNILVLTDQDVELGS 60
QY 61 MOVNKTTRIMEGGTHFTNAFVTPMCCPSRSSILTGKYVHNHTYTNNECSSPSWQA 120
DB 61 MOVNKTTRIMEGGHAFFNAFVTPMCCPSRSSILTGKYVHNHTYTNNECSSPSWQA 120
QY 121 QHESRTFAVYLNSTGYRTAFFGKYLNEYGSYVPPGKWEVGLLKNRSFYNTLCRNGVK 180
DB 121 QHESRTFAVYLNSTGYRTAFFGKYLNEYGSYVPPGKWEVGLLKNRSFYNTLCRNGVK 180
QY 181 EXHGSYSDYLDLITNDVSFFRTSKOYHPHVPVLAIVSHAAHPGSDSAPQYSLRFP 240
DB 181 EXHGSYSDYLDLITNDVSFFRTSKOYHPHVPVLAIVSHAAHPGSDSAPQYSLRFP 240
QY 241 NASOHLTPSYNAPNDKQWIMRYTGPMPKPIHMEFTNMLQRKLOTLMASVDSMETIYNM 300
DB 241 NASOHLTPSYNAPNDKQWIMRYTGPMPKPIHMEFTNMLQRKLOTLMASVDSMETIYNM 300
QY 301 LVETGELDNITYIYVADHGVHIGQFGLVGKSNPVEFDIRVPPYVRGPNVEAGCLNPHIV 360
DB 301 LVETGELDNITYIYVADHGVHIGQFGLVGKSNPVEFDIRVPPYVRGPNVEAGCLNPHIV 360
QY 361 LNTDLAPTILDIAGLIDIPADMDGKSLKLLDTERPVRNPHLKKRVRWDSFLVERGKLL 420
DB 361 LNTDLAPTILDIAGLIDIPADMDGKSLKLLDTERPVRNPHLKKRVRWDSFLVERGKLL 420

QY 421 HKRDNDKVDQAQENFLPKYQRYKDLQRAEYQTAQELGQKQWQCVEDATGKLKHKCKGP 480
DB 421 HKRDNDKVDQAQENFLPKYQRYKDLQRAEYQTAQELGQKQWQCVEDATGKLKHKCKGP 480
QY 481 MRLGSSRLSNLVPKYGGSEACTCDSDGYKLSLAGRRKKLFKKKYKASYVRSRSIRSV 540
DB 481 MRLGSSRLSNLVPKYGGSEACTCDSDGYKLSLAGRRKKLFKKKYKASYVRSRSIRSV 540
QY 541 AIEVDGRVYHVGDAAPRNLTQKHWPGAPEDQDDKGGDFSGTGGGLPDYSAANPIKVT 600
DB 541 AIEVDGRVYHVGDAAPRNLTQKHWPGAPEDQDDKGGDFSGTGGGLPDYSAANPIKVT 600
QY 601 HRCYILENTVQCDLDLYKSLQAMKDKHLHDHEIETLQNKIKNLREVGRHLKKCRPEEC 660
DB 601 HRCYILENTVQCDLDLYKSLQAMKDKHLHDHEIETLQNKIKNLREVGRHLKKCRPEEC 660
QY 661 DCHKISYHTQHGRLKGRSSSLHPRKGLQEKQVWLLREQKQKRLKRLQNDTC 720
DB 661 DCHKISYHTQHGRLKGRSSSLHPRKGLQEKQVWLLREQKQKRLKRLQNDTC 720
QY 721 SMPGLTCFTHDQHWQTAPFWTLGPFCACTSANNTYWCRTINETHNLFCEFATGFLE 780
DB 721 SMPGLTCFTHDQHWQTAPFWTLGPFCACTSANNTYWCRTINETHNLFCEFATGFLE 780
QY 781 YFDLNTDPPYQLANVNTLDRVNLQHLVQLMELRSCKGYKQCNPRTRNMDLGLKGGSYE 840
DB 781 YFDLNTDPPYQLANVNTLDRVNLQHLVQLMELRSCKGYKQCNPRTRNMDLGLKGGSYE 840
QY 841 QYRQFORRWKPEKRPSSKSLGOLWEGWEG 870
DB 841 QYRQFORRWKPEKRPSSKSLGOLWEGWEG 870
RESULT 11
AAV66648
ID AAV66648 standard; protein; 867 AA.
XX
AC AAV66648;
XX
DT 05-APR-2000 (first entry)
XX
DE Membrane-bound protein PRO1120.
XX
KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
KW pharmaceutical; receptor immunoadhesin; gene mapping.
XX
OS Homo sapiens.
XX
PN WO9963088-A2.
XX
PD 09-DEC-1999.
XX
PF 02-JUN-1999; 99WO-US12252.
XX
PR 02-JUN-1998; 98US-0087607.
PR 02-JUN-1998; 98US-0087609.
PR 02-JUN-1998; 98US-0087759.
PR 03-JUN-1998; 98US-0087827.
PR 04-JUN-1998; 98US-0088021.
PR 04-JUN-1998; 98US-0088025.
PR 04-JUN-1998; 98US-0088028.
PR 04-JUN-1998; 98US-0088029.
PR 04-JUN-1998; 98US-0088030.
PR 04-JUN-1998; 98US-0088033.
PR 04-JUN-1998; 98US-0088326.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088212.
PR 05-JUN-1998; 98US-0088217.
PR 09-JUN-1998; 98US-0088655.
PR 10-JUN-1998; 98US-0088722.
PR 10-JUN-1998; 98US-0088730.

PR 10-JUN-1998; 98US-0088734.
PR 10-JUN-1998; 98US-0088738.
PR 10-JUN-1998; 98US-0088740.
PR 10-JUN-1998; 98US-0088741.
PR 10-JUN-1998; 98US-0088742.
PR 10-JUN-1998; 98US-0088810.
PR 10-JUN-1998; 98US-0088811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 10-JUN-1998; 98US-0088826.
PR 11-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089090.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.
PR 16-JUN-1998; 98US-0089512.
PR 16-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089532.
PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089598.
PR 17-JUN-1998; 98US-0089599.
PR 17-JUN-1998; 98US-0089600.
PR 17-JUN-1998; 98US-0089653.
PR 18-JUN-1998; 98US-0089801.
PR 18-JUN-1998; 98US-0089907.
PR 18-JUN-1998; 98US-0089908.
PR 19-JUN-1998; 98US-0089947.
PR 19-JUN-1998; 98US-0089948.
PR 19-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 24-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 24-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090862.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 09-JUL-1998; 98US-0091982.
PR 10-JUL-1998; 98US-0092182.
PR 20-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.

PR 30-JUL-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095308.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.
PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 18-AUG-1998; 98US-0097022.
PR 19-AUG-1998; 98US-0097141.
PR 20-AUG-1998; 98US-0097218.
PR 24-AUG-1998; 98US-0097661.
PR 26-AUG-1998; 98US-0097951.
PR 26-AUG-1998; 98US-0097952.
PR 26-AUG-1998; 98US-0097954.
PR 26-AUG-1998; 98US-0097955.
PR 26-AUG-1998; 98US-0097971.
PR 26-AUG-1998; 98US-0097974.
PR 26-AUG-1998; 98US-0097978.
PR 26-AUG-1998; 98US-0097979.
PR 26-AUG-1998; 98US-0097986.
PR 26-AUG-1998; 98US-0098014.
PR 31-AUG-1998; 98US-0098525.
PR 16-SEP-1998; 98US-0100634.
PR 12-JAN-1999; 99US-0115565.
XX
PA (GETH) GENENTECH INC.
PI Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
PI Wood WI, Yuan J;
XX
XX WPI: 2000-072883/06.
XX N-PSDB; AAZ64969.
XX Membrane-bound proteins and related nucleotide sequences -
XX claim 12; Fig 47; 822pp; English.
XX
XX The invention provides membrane-bound PRO polypeptides and
XX polynucleotides encoding them. The PRO sequences of the invention were
XX identified based on extracellular domain homology screening. The PRO
XX sequences have homology with proteins including LDL receptors, TIE
XX ligands and various enzymes. The membrane-bound proteins and receptor
XX molecules are useful as pharmaceutical and diagnostic agents. Receptor
XX immunoadhesins, for instance, can be used as therapeutic agents to block
XX receptor-ligand interactions. The membrane-bound proteins can also be
XX employed for screening of potential peptide or small molecule inhibitors
XX of the relevant receptor/ligand interaction. The PRO encoding sequences
XX are useful as hybridization probes, in chromosome and gene mapping and in
XX the generation of antisense RNA and DNA. PRO nucleic acid sequences
XX will also be useful for the preparation of PRO polypeptides, especially
XX by recombinant techniques.

SQ	Sequence	867 AA;
	Query Match	99.3%; Score 4719.5; DB 21; Length 867;
	Best Local Similarity	99.5%; Pred. No. 0;
	Matches 866; Conservative	0; Mismatches 1; Indels 3; Gaps 1;
QY	1	MGPPSLVLCLLSATVFSLLGSSAFSLSHRLKGRFORDRNRIPNIIILVLTDDQVLELGS 60
DB	1	MGPPSLVLCLLSATVFSLLGSSAFSLSHRLKGRFORDRNRIPNIIILVLTDDQVLELGS 60
QY	61	MOVNKKTRRIMEQGGTHIFNAFTVTPMCCPSRSILITGKYVHNHTYTNNECSPSQA 120
DB	61	MOVNKKTRRIMEQGGTHIFNAFTVTPMCCPSRSILITGKYVHNHTYTNNECSPSQA 120
QY	121	QHSRTFAVLNSTGYRTAFGKYLNEYSGYVPPGKEMVGLLKNRFPYNTLCRNGVK 180
DB	121	QHSRTFAVLNSTGYRTAFGKYLNEYSGYVPPGKEMVGLLKNRFPYNTLCRNGVK 180
QY	181	EKGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPQYSLRFP 240
DB	181	EKGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAPHGPDSPAPQYSLRFP 240
QY	241	NASQHTPSYNAAPNDKHWIMYTGPMKPIHMEFTNMLQKELQTLMSVDDSMETIYNN 300
DB	241	NASQHTPSYNAAPNDKHWIMYTGPMKPIHMEFTNMLQKELQTLMSVDDSMETIYNN 300
QY	301	LVETGELDNTYIVYTADHGVIHQFGLVKGKSNPYEFDIRVPPYVGRPNVAGCLNPHIV 360
DB	301	LVETGELDNTYIVYTADHGVIHQFGLVKGKSNPYEFDIRVPPYVGRPNVAGCLNPHIV 360
QY	361	LNIDLAPTILDIAGLDIPADMDGKSILKLDLTERPVNRPHLKKOMVWDSFLVERGKLL 420
DB	361	LNIDLAPTILDIAGLDIPADMDGKSILKLDLTERPVNRPHLKKOMVWDSFLVERGKLL 420
QY	421	HKRNDKVDQAEENFLPKYQKVDLQRAEYQACQLQKQKQCVEDATGKLLHKCKGP 480
DB	421	HKRNDKVDQAEENFLPKYQKVDLQRAEYQACQLQKQKQCVEDATGKLLHKCKGP 480
QY	481	MRLGSRALSNLVPKYGGQSEACTCDSDYKLSLAGRBKLPKKYKASYRSTRSISV 540
DB	481	MRLGSRALSNLVPKYGGQSEACTCDSDYKLSLAGRBKLPKKYKASYRSTRSISV 540
QY	541	AIEDVGRVYHVLGDAAPRNLYKRWHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKYT 600
DB	541	AIEDVGRVYHVLGDAAPRNLYKRWHPGAPEDQDDKGGDFSGTGGLPDYSAANPIKYT 600
QY	601	HRCVILENDTVQCDLDLYKSIQAKWKHLHIHIEITLQNKIKNLRVFGHLKKRPEEC 660
DB	601	HRCVILENDTVQCDLDLYKSIQAKWKHLHIHIEITLQNKIKNLRVFGHLKKRPEEC 660
QY	661	DCHKISYHTQKGRKLRHSGSLHPPFRKGLQEKQKWLRLREQKPKKLRKLLKRLQNNDT 720
DB	661	DCHKISYHTQKGRKLRHSGSLHPPFRKGLQEKQKWLRLREQKPKKLRKLLKRLQNNDT 720
QY	721	SMPLGLTCFTHDNQHWOTAPFTLGPCCACTSANNTYWCRTINETHNFLECEPATGELE 780
DB	721	SMPLGLTCFTHDNQHWOTAPFTLGPCCACTSANNTYWCRTINETHNFLECEPATGELE 780
QY	781	YFDLNTDYPQLANAVNTLDRLVNLQHLVQLMELSCKYKQCNPRTRNMDLGLKDGGSYE 840
DB	781	YFDLNTDYPQLANAVNTLDRLVNLQHLVQLMELSCKYKQCNPRTRNMDLGLKDGGSYE 840
QY	841	QYRQQRKKNPKRPSKSLQLMEWGEG 870
DB	838	QYRQQRKKNPKRPSKSLQLMEWGEG 867

RESULT 12

ID AAU12181 standard; Protein; 867 AA.

XX AAU12181;

XX

DT	24-OCT-2001	(first entry)
XX	Human PRO1120 polypeptide sequence.	
XX	Human secretory and transmembrane; PRO; mammalian; cancer; lung;	
KW	breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;	
KW	cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;	
KW	adipocyte; A-peptide; factor VIIA; gene therapy.	
OS	Homo sapiens.	
XX	WO200140466-A2.	
XX	07-JUN-2001.	
XX	01-DEC-2000; 2000WO-US32678.	
XX	01-DEC-1999; 99WO-US28301.	
PR	01-DEC-1999; 99WO-US28634.	
PR	02-DEC-1999; 99WO-US28551.	
PR	02-DEC-1999; 99WO-US28564.	
PR	02-DEC-1999; 99WO-US28565.	
PR	09-DEC-1999; 99US-0170262.	
PR	16-DEC-1999; 99WO-US30095.	
PR	20-DEC-1999; 99WO-US30911.	
PR	20-DEC-1999; 99WO-US30999.	
PR	30-DEC-1999; 99WO-US31243.	
PR	06-JAN-2000; 2000WO-US00277.	
PR	06-JAN-2000; 2000WO-US00376.	
PR	11-FEB-2000; 2000WO-US03565.	
PR	18-FEB-2000; 2000WO-US04341.	
PR	18-FEB-2000; 2000WO-US04342.	
PR	22-FEB-2000; 2000WO-US04414.	
PR	24-FEB-2000; 2000WO-US04914.	
PR	24-FEB-2000; 2000WO-US05004.	
PR	01-MAR-2000; 2000WO-US05601.	
PR	20-MAR-2000; 2000WO-US07377.	
PR	21-MAR-2000; 2000WO-US07532.	
PR	30-MAR-2000; 2000WO-US08439.	
PR	17-MAY-2000; 2000WO-US13705.	
PR	22-MAY-2000; 2000WO-US14042.	
PR	30-MAY-2000; 2000WO-US14941.	
PR	02-JUN-2000; 2000WO-US15264.	
PR	10-NOV-2000; 2000WO-US30873.	
XX	{GETH } GENENTECH INC.	
XX	Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;	
PI	Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;	
PI	Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;	
XX	WPI; 2001-408281/43.	
DR	N-PSDB; AAS21253.	
XX	Isolated, secretory and transmembrane PRO polypeptide used to detect	
PT	other PRO polypeptides, link bioactive molecules to cells expressing	
PT	PRO polypeptides, and detect the presence of mammalian tumours e.g.	
PT	lung, breast, prostate, cervical	
XX	Claim 12; Fig 20; 813pp; English.	
XX	AAU12172-AAU12446 represent novel human secretory and transmembrane	
CC	PRO polypeptides. The PRO polypeptides are useful to detect other	
CC	PRO polypeptides, to link bioactive molecules to cells expressing	
CC	PRO polypeptides, to modulate biological activities of cells expressing	
CC	PRO polypeptides, and to detect the presence of mammalian lung, colon,	
CC	breast, prostate, rectal, cervical or liver tumours by comparing PRO	
CC	polypeptide expression in a cell sample to that in a control sample.	
CC	Some of the 275 sequences are also useful to stimulate the release of	
CC	tumour necrosis factor-alpha (TNF-alpha) from human blood, the	
CC	proliferation or differentiation of chondrocytes, the proliferation or	
CC	gene expression in pericyte cells, the release of proteoglycans from	
CC	cartilage, the proliferation of inner ear utricular supporting cells or	

CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 XX
 SQ Sequence 867 AA;
 Query Match 99.3%; Score 4719.5; DB 22; Length 867;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
 QY 1 MGPPSLVCLLSATVFLSGSSAFSLSHRLKGRFQDRNIRPNILVLTDDQDELGS 60
 DB 1 MGPPSLVCLLSATVFLSGSSAFSLSHRLKGRFQDRNIRPNILVLTDDQDELGS 60
 QY 61 MOVNKKTRRIMEQGTTHFNAFVTPMCCPSRSSILTKGVVHNHTYNNENCSPSMQA 120
 DB 61 MOVNKKTRRIMEQGAHFNAFVTPMCCPSRSSILTKGVVHNHTYNNENCSPSMQA 120
 QY 121 QHESRTFAVLNLTGYRTAFPGKYLNEVNGSYVPPGKWKVGLLQNSRFNYTLCRNGVK 180
 DB 121 QHESRTFAVLNLTGYRTAFPGKYLNEVNGSYVPPGKWKVGLLQNSRFNYTLCRNGVK 180
 QY 181 EKHGSDYSKDYLTDLTNDVSVSFRTSKOMYPHPVLMVISHAAPHGPDSPAYSRLLFP 240
 DB 181 EKHGSDYSKDYLTDLTNDVSVSFRTSKOMYPHPVLMVISHAAPHGPDSPAYSRLLFP 240
 QY 241 NASQHITPSYVAPNPKHIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
 DB 241 NASQHITPSYVAPNPKHIMRYTGPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN 300
 QY 301 LVETGELDNTYIVYVADHGYHIGQGLVKGSMPEYEDIRVPFYVGPVNYEAGCLNPHIV 360
 DB 301 LVETGELDNTYIVYVADHGYHIGQGLVKGSMPEYEDIRVPFYVGPVNYEAGCLNPHIV 360
 QY 361 LNIDLAPTLIDLAGLIDPADMDGKSLKLLDTERPVNRPHLKKQVRWRDSFLVERGKLL 420
 DB 361 LNIDLAPTLIDLAGLIDPADMDGKSLKLLDTERPVNRPHLKKQVRWRDSFLVERGKLL 420
 QY 421 HKRNDKVDQAENFLPKYQVKDLQRAEYQTAQCEQLGQKQVCEVATGKLLHKCKGP 480
 DB 421 HKRNDKVDQAENFLPKYQVKDLQRAEYQTAQCEQLGQKQVCEVATGKLLHKCKGP 480
 QY 481 MRLGSSRALSLVPKYQGGSEACTDSGDYKLSLAGRRKLPKKTAKSYVRSRSIRSV 540
 DB 481 MRLGSSRALSLVPKYQGGSEACTDSGDYKLSLAGRRKLPKKTAKSYVRSRSIRSV 540
 QY 541 ALEVDRVYHVLGDAQAPRLTKRWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
 DB 541 ALEVDRVYHVLGDAQAPRLTKRWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
 QY 601 HRCYILENDTVQCDLDLYKSLQAKWKHLHDHIEITLQNKIKWLREVRGHLKKGPEEC 660
 DB 601 HRCYILENDTVQCDLDLYKSLQAKWKHLHDHIEITLQNKIKWLREVRGHLKKGPEEC 660
 QY 661 DCHKLSYTHQKRLKRGSSLHPKRGKQKDKVLLREOKKKLKLRLQNNDT 720
 DB 661 DCHKLSYTHQKRLKRGSSLHPKRGKQKDKVLLREOKKKLKLRLQNNDT 720
 QY 721 SMPGLTCTFDHNDHQWQTAPFWTLGLPFCACTSANNNYTWCMRTINETNFIKCEPATGFL 780
 DB 721 SMPGLTCTFDHNDHQWQTAPFWTLGLPFCACTSANNNYTWCMRTINETNFIKCEPATGFL 780
 QY 781 YFDLNTDYPQLMNAVNTLDRVLNQLHVLMLSCSKYKQCNPRTRNMDLGLKDGSSYE 840
 DB 781 YFDLNTDYPQLMNAVNTLDRVLNQLHVLMLSCSKYKQCNPRTRNMDLGLKDGSSYE 840
 QY 841 QYRQFORRWKPEMKRPSKSLGQLMGEGWEG 870

DB 838 QYRQFORRWKPEMKRPSKSLGQLMGEGWEG 867
 RESULT 13
 AAB65171
 ID AAB65171 standard; Protein; 867 AA.
 XX AC AAB65171;
 XX DT 02-APR-2001 (first entry)
 XX DE Human PRO1120 (UNQ559) protein sequence SEQ ID NO:84.
 XX KW Human; secreted and transmembrane protein; PRO; cytostatic;
 KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
 KW diagnostic assay.
 XX OS Homo sapiens.
 XX PN WO2000073454-A1.
 XX PD 07-DEC-2000.
 XX PF 30-MAR-2000; 2000WO-US08439.
 XX PR 02-JUN-1999; 99WO-US12252.
 XX PR 23-JUN-1999; 99US-0141037.
 XX PR 07-JUL-1999; 99US-0143048.
 XX PR 20-JUL-1999; 99US-0144758.
 XX PR 26-JUL-1999; 99US-0145698.
 XX PR 28-JUL-1999; 99US-0146222.
 XX PR 17-AUG-1999; 99US-0149396.
 XX PR 15-SEP-1999; 99WO-US21090.
 XX PR 15-SEP-1999; 99WO-US21547.
 XX PR 08-OCT-1999; 99US-0158663.
 XX PR 30-NOV-1999; 99WO-US28313.
 XX PR 01-DEC-1999; 99WO-US28301.
 XX PR 16-DEC-1999; 99WO-US30095.
 XX PR 20-DEC-1999; 99WO-US30911.
 XX PR 05-JAN-2000; 2000WO-US00219.
 XX PR 06-JAN-2000; 2000WO-US00376.
 XX PR 11-FEB-2000; 2000WO-US03565.
 XX PR 18-FEB-2000; 2000WO-US04341.
 XX PR 22-FEB-2000; 2000WO-US04414.
 XX PR 24-FEB-2000; 2000WO-US04914.
 XX PR 02-MAR-2000; 2000WO-US05841.
 XX PR 15-MAR-2000; 2000WO-US06884.
 XX PR 20-MAR-2000; 2000WO-US07377.
 (GETH) GENENTECH INC.
 XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
 PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
 PI Grimaldi CJ, Gurney AJ, Kljavin IJ, Napier MA, Pan J, Paoni NF;
 PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
 PI Zhang Z;
 XX WPI; 2001-032160/04.
 DR N-PSDB; AAF44115.
 XX PRO polynucleotides used to produce polypeptides used to target
 PT bioactive molecules such as toxins, radiolabels or antibodies, to
 PT specific cells, to cause targeted cell death -
 XX Claim 12; Fig 47; 935pp; English.
 CC The present invention describes human secreted and transmembrane PRO
 CC proteins. The PRO proteins have cytostatic activity. The PRO proteins
 CC can be used for targeted delivery of bioactive molecules, such as
 CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
 CC sequences, and their fragments, can be used as hybridization probes, in
 CC chromosomal and gene mapping, and in the generation of anti-sense RNA

CC and DNA. They may also be used to produce transgenic animals which are
 CC used to develop and screen therapeutically useful reagents. The PRO
 CC nucleotide and protein sequence can be used for tissue typing and in
 CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
 CC AAF44270 to AAF4470 represent PCR primers and hybridisation probes used
 CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
 CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
 CC sequences given in the exemplification of the present invention.

XX Sequence 867 AA;

Query Match	99.3%	Score	4719.5;	DB	22;	Length	867;
Best Local Similarity	99.5%	Pred. No.	0;				
Matches	866;	Conservative	0;	Mismatches	1;	Indels	3;
						Gaps	1;

Qy	1	MGPPSLVCLLSATVPSLLGGSAFLSHRLKGRFORDRNRIPNIIILVLTDDQDVELGS	60
Db	1	MGPPSLVCLLSATVPSLLGGSAFLSHRLKGRFORDRNRIPNIIILVLTDDQDVELGS	60
Qy	61	MOVNKTTRIMEGGGTHFINAFVTTMCCPSRSSILTGKYYVHNNTYNNENCSPPSWQA	120
Db	61	MOVNKTTRIMEGGGTHFINAFVTTMCCPSRSSILTGKYYVHNNTYNNENCSPPSWQA	120
Qy	121	QHSRTFAVYLNSTGYRTAFYGGVLYNEVNGSVYPPGKMEVGLLNKNSRFYNTLCRNGVK	180
Db	121	QHSRTFAVYLNSTGYRTAFYGGVLYNEVNGSVYPPGKMEVGLLNKNSRFYNTLCRNGVK	180
Qy	181	EKHGSDYSKDYLTDLITNDVSFFRTSKKYPHRLVLMVISHAAPHGPDSPAPYSRLFP	240
Db	181	EKHGSDYSKDYLTDLITNDVSFFRTSKKYPHRLVLMVISHAAPHGPDSPAPYSRLFP	240
Qy	241	NASCHITPSNYAPNPKGIMIEYTPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN	300
Db	241	NASCHITPSNYAPNPKGIMIEYTPMKPIHMEFTNMLQKRLQTLMSVDDSMETIYNN	300
Qy	301	LVTGELDNYYIYVYADHGYHICQGLVKGKSNPYEFDIRVPPYVGRPNVAGCLNPHIV	360
Db	301	LVTGELDNYYIYVYADHGYHICQGLVKGKSNPYEFDIRVPPYVGRPNVAGCLNPHIV	360
Qy	361	LNLDLAPTLIDAGLIDIPADMCKSLIKLLDTERPVNREHAKKRVWRDSEFLVERGKLL	420
Db	361	LNLDLAPTLIDAGLIDIPADMCKSLIKLLDTERPVNREHAKKRVWRDSEFLVERGKLL	420
Qy	421	HKRDNDKVDAAQENFLPKYQVXDLQCORAEYQACEQLGQKQCVEDATGKLKHKCKGP	480
Db	421	HKRDNDKVDAAQENFLPKYQVXDLQCORAEYQACEQLGQKQCVEDATGKLKHKCKGP	480
Qy	481	MRLLGGSRLSNLVPKYGGGSEACTCSDGYKLSLAGREKLLFKKKYKASYRSHSISV	540
Db	481	MRLLGGSRLSNLVPKYGGGSEACTCSDGYKLSLAGREKLLFKKKYKASYRSHSISV	540
Qy	541	AIEVDGRVYHVLGDAAPENLTKRHWPGAPEDQDDKDGDFSGTGGLPDYSAANPIKVT	600
Db	541	AIEVDGRVYHVLGDAAPENLTKRHWPGAPEDQDDKDGDFSGTGGLPDYSAANPIKVT	600
Qy	601	HRCVILENDTVQCDLIDYKSLQAWKHKLHIEITLQNKIKNIREVGRGHKKRPEC	660
Db	601	HRCVILENDTVQCDLIDYKSLQAWKHKLHIEITLQNKIKNIREVGRGHKKRPEC	660
Qy	661	DCHKISVHTQKGRKLRHSGSLHPPFKGLQEKDKVWLLREQKRLKRLKRLQNNDTG	720
Db	661	DCHKISVHTQKGRKLRHSGSLHPPFKGLQEKDKVWLLREQKRLKRLKRLQNNDTG	720
Qy	721	SMPLGTCFTHDNQHWTAPFTWTLGPFCACTSANNTYWCRTINETHNPLFCEFATGPLE	780
Db	721	SMPLGTCFTHDNQHWTAPFTWTLGPFCACTSANNTYWCRTINETHNPLFCEFATGPLE	780
Qy	781	YFDLNTDPYOLMNAVNTLDRDVLNQLHVLQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE	840
Db	781	YFDLNTDPYOLMNAVNTLDRDVLNQLHVLQVLMELRSCKYKQCNPRTRNMDLGLKDGGSYE	840
Qy	841	QYQFORRKWPKMKRPSKSLGQLWEGWEG	870
Db	841	QYQFORRKWPKMKRPSKSLGQLWEGWEG	870

Db 838 QYQFORRKWPKMKRPSKSLGQLWEGWEG 867

RESULT 14

AB895459

ID ABB95459 standard; Protein; 867 AA.

XX AC ABB95459;

XX AC ABB95459;

XX AC ABB95459;

DT 19-JUL-2002 (first entry)

XX DE Human angiogenesis related protein PRO120 SEQ ID NO: 74.

XX KW Human; angiogenesis; PRO protein; cardiovascularisation; wound; cancer;
 KW atherosclerosis; cardiac hypertrophy; gene therapy; endothelial disorder;
 KW cardiatic; cytostatic; antiangiogenic; hypotensive; vulnerary;
 KW antiarterioeclerotic.

XX OS Homo sapiens.

XX PN WO200208284-A2.

XX PD 31-JAN-2002.

XX PF 09-JUL-2001; 2001WO-US21735.

XX PR 20-JUL-2000; 2000US-219556P.

XX PR 25-JUL-2000; 2000US-220624P.

XX PR 25-JUL-2000; 2000US-220664P.

XX PR 28-JUL-2000; 2000WO-US20710.

XX PR 02-AUG-2000; 2000US-222695P.

XX PR 17-AUG-2000; 2000US-0643657.

XX PR 23-AUG-2000; 2000WO-US23522.

XX PR 24-AUG-2000; 2000WO-US23328.

XX PR 07-SEP-2000; 2000US-230378P.

XX PR 15-SEP-2000; 2000US-000000P.

XX PR 18-SEP-2000; 2000US-0664610.

XX PR 18-SEP-2000; 2000US-0665350.

XX PR 24-OCT-2000; 2000US-242922P.

XX PR 08-NOV-2000; 2000US-0709238.

XX PR 08-NOV-2000; 2000WO-US30952.

XX PR 10-NOV-2000; 2000WO-US30873.

XX PR 01-DEC-2000; 2000WO-US32678.

XX PR 20-DEC-2000; 2000US-0747259.

XX PR 20-DEC-2000; 2000WO-US34956.

XX PR 22-JAN-2001; 2001US-0767609.

XX PR 28-FEB-2001; 2001US-0796498.

XX PR 28-FEB-2001; 2001WO-US05520.

XX PR 01-MAR-2001; 2001WO-US06666.

XX PR 09-MAR-2001; 2001US-0802706.

XX PR 14-MAR-2001; 2001US-0808689.

XX PR 22-MAR-2001; 2001US-0816744.

XX PR 05-APR-2001; 2001US-0828366.

XX PR 10-MAY-2001; 2001US-0854208.

XX PR 10-MAY-2001; 2001US-0854280.

XX PR 23-MAY-2001; 2001US-0866028.

XX PR 25-MAY-2001; 2001US-0866034.

XX PR 25-MAY-2001; 2001WO-US17092.

XX PR 30-MAY-2001; 2001US-0870574.

XX PR 30-MAY-2001; 2001WO-US17443.

XX PR 01-JUN-2001; 2001WO-US17800.

XX PR 20-JUN-2001; 2001WO-US19692.

XX PR 28-JUN-2001; 2001WO-US00000.

XX (GETH) GENENTECH INC.

PA (BAKE/) BAKER K P.

PA (FERE/) FERRARA N.

PA (GERB/) GERBER H.

PA (GERE/) GERRITSEN M E.

PA (GODD/) GODDARD A.

PA (GODO/) GODOWSKI P J.

PA (GURN/) GURNEY A L.

PA (HILL/) HILLAN K J.

(MARS/) MARSTERS S A.
 PA (PANJ/) PAN J.
 PA (PAON/) PAONI N F.
 PA (STEP/) STEPHAN J F.
 PA (WATA/) WATANABE C K.
 PA (WILL/) WILLIAMS F M.
 PA (WOOD/) WOOD W I.
 XX Baker KP, Ferrata N, Gerber H, Gerritsen ME, Goddard A;
 PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF;
 PI Stephan JF, Watanabe CK, Williams FM, Wood WI, Ye W;
 XX WPI; 2002-171999/22.
 DR N-PSDB; ABU95597.
 XX
 PT One hundred and eighty seven nucleic acids encoding PRO polypeptides,
 PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial
 PT infarction), endothelial or angiogenic disorders in a mammal -
 XX
 PS Claim 11; Fig 74; 567pp; English.
 XX
 CC The present invention provides the protein and coding sequences of human
 CC PRO proteins. These are useful for treating or diagnosing a
 CC cardiovascular, endothelial or angiogenic disorder, including cardiac
 CC hypertrophy, trauma, cancer, age-related macular degeneration,
 CC atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis,
 CC angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour
 CC angiogenesis (such as breast carcinoma and liver carcinoma) and wound
 CC healing. The present sequence is a PRO protein of the invention.
 XX
 XX Sequence 867 AA;
 SQ

Query Match 99.3%; Score 4719.5; DB 23; Length 867;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
 1 MGPPSLVLCCLLSATVFSLLGGSAFLSHRLKGRFORRRIRNIIILVLTDDQVGLS 60
 1 MGPPSLVLCCLLSATVFSLLGGSAFLSHRLKGRFORRRIRNIIILVLTDDQVGLS 60
 61 MQVMNKTREIMEQGGHFIINAFVTPMCCPSRSSILTKYVHNHNTYNNENCSPSWQA 120
 61 MQVMNKTREIMEQGGHFIINAFVTPMCCPSRSSILTKYVHNHNTYNNENCSPSWQA 120
 121 QHESRTFAVILNSTGYRTAFPGKYLNENYGVVPPGKWKVGLLKNRPFYNTLCRNGVK 180
 121 QHESRTFAVILNSTGYRTAFPGKYLNENYGVVPPGKWKVGLLKNRPFYNTLCRNGVK 180
 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKOMYPHRRPVLAVISHAAPHGPDSPAPYSLFP 240
 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKOMYPHRRPVLAVISHAAPHGPDSPAPYSLFP 240
 241 NASQHTTPSYNAPNDKHIMRYTGPMKPIMHETNMLQKGLTLMVDSDMETIYNN 300
 241 NASQHTTPSYNAPNDKHIMRYTGPMKPIMHETNMLQKGLTLMVDSDMETIYNN 300
 301 LVETGELDNTIYVYADHGYHIGQGLVKGKSMPEYFDIRVFPVYVGNVVEAGCLNPHIV 360
 301 LVETGELDNTIYVYADHGYHIGQGLVKGKSMPEYFDIRVFPVYVGNVVEAGCLNPHIV 360
 361 LNIDLAFTILDIAGLDIPADMGSKSILKLLDTERPVNRFHLKKQKVRWDSFLVERGKLL 420
 361 LNIDLAFTILDIAGLDIPADMGSKSILKLLDTERPVNRFHLKKQKVRWDSFLVERGKLL 420
 421 HKRDNDKVAQENFUPKYQVRKDLQRAEYQTAQEQLGKQKQCVEDATGKLKHKCKGP 480
 421 HKRDNDKVAQENFUPKYQVRKDLQRAEYQTAQEQLGKQKQCVEDATGKLKHKCKGP 480
 481 MRLGGSRLSNLVPKYGGGSEACTCDSDGYKLSLAGRRKKLFKKKYKASVYRSRSRSV 540
 481 MRLGGSRLSNLVPKYGGGSEACTCDSDGYKLSLAGRRKKLFKKKYKASVYRSRSRSV 540
 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGLPDYSAANPIKVT 600

Db 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGLPDYSAANPIKVT 600
 Qy 601 HRCYILENDTVQCDLDLYKSLQAMKDHKLHIDHEITLQNKIKNLREVRGHLKKRPEEC 660
 Db 601 HRCYILENDTVQCDLDLYKSLQAMKDHKLHIDHEITLQNKIKNLREVRGHLKKRPEEC 660
 Qy 661 DCHKISYHTQHKGRKLRHSGSSSLHFFRKGLOEKQKVLLREQKRKKLRLKRLQNNDTG 720
 Db 661 DCHKISYHTQHKGRKLRHSGSSSLHFFRKGLOEKQKVLLREQKRKKLRLKRLQNNDTG 720
 Qy 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFL 780
 Db 721 SMPGLTCFTHDNQHWOTAPFWTLGPFCACTSANNTYWCMTINETHNFCEPATGFL 780
 Qy 781 YFDLNTDPYQLMNAVNTLDRDLNQLHVLQVLMELRSCKYKQCNPRTRNMOLGLKDGGSYE 840
 Db 781 YFDLNTDPYQLMNAVNTLDRDLNQLHVLQVLMELRSCKYKQCNPRTRNMOLGLKDGGSYE 840
 Qy 841 QYRQFQRKRWPEMKRPPSSKSLGQLWEGWEG 870
 Db 838 QYRQFQRKRWPEMKRPPSSKSLGQLWEGWEG 867
 RESULT 15
 ABB84853
 ID ABB84853 standard; Protein; 867 AA.
 XX ABB84853;
 XX
 DT 16-MAY-2002 (first entry)
 XX
 DE Human PRO1120 protein sequence SEQ ID NO:74.
 XX
 KW Human; angiogenesis; cardiant; cytostatic; antiangiogenic; hypotensive;
 KW vulnerary; antiarteriosclerotic; PRO agonist; PRO antagonist; trauma;
 KW gene therapy; cardiovascular disorder; endothelial disorder; cancer;
 KW angiogenic disorder; cardiac hypertrophy; atherosclerosis; hypertension;
 KW age-related macular degeneration; arterial restenosis; angina;
 KW rheumatoid arthritis; myocardial infarction; thrombophlebitis;
 KW lymphangitis; tumour angiogenesis; breast carcinoma; liver carcinoma;
 KW wound healing; chromosome mapping; gene mapping.
 XX
 OS Homo sapiens.
 XX
 PN WO200200690-A2.
 XX
 XX 03-JAN-2002.
 PD
 PF 20-JUN-2001; 2001WO-US19692.
 PF
 PF 23-JUN-2000; 2000US-213637P.
 PR 20-JUL-2000; 2000US-219556P.
 PR 25-JUL-2000; 2000US-220824P.
 PR 25-JUL-2000; 2000US-220664P.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 02-AUG-2000; 2000US-222695P.
 PR 17-AUG-2000; 2000US-0643657.
 PR 23-AUG-2000; 2000WO-US23522.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 07-SEP-2000; 2000US-230978P.
 PR 18-SEP-2000; 2000US-0664610.
 PR 18-SEP-2000; 2000US-0665350.
 PR 24-OCT-2000; 2000US-242922P.
 PR 08-NOV-2000; 2000US-0709238.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000US-0747259.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 22-JAN-2001; 2001US-0767609.
 PR 28-FEB-2001; 2001US-0796498.
 PR 28-FEB-2001; 2001WO-US06520.

PR 01-MAR-2001; 2001WO-US06666.
 PR 09-MAR-2001; 2001US-0802706.
 PR 14-MAR-2001; 2001US-0809689.
 PR 22-MAR-2001; 2001US-0816744.
 PR 05-APR-2001; 2001US-0828366.
 PR 10-MAY-2001; 2001US-0854208.
 PR 10-MAY-2001; 2001US-0854280.
 PR 25-MAY-2001; 2001US-0866028.
 PR 25-MAY-2001; 2001US-0866034.
 PR 25-MAY-2001; 2001WO-US17092.
 PR 30-MAY-2001; 2001US-0870574.
 PR 30-MAY-2001; 2001WO-US17443.
 PR 01-JUN-2001; 2001WO-US17800.
 XX (GETH) GENENTECH INC.
 XX
 XX Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A; Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF; PI Stephan JF, Watanabe CK, Williams PW, Wood WL, Ye W;
 XX
 DR MPI; 2002-090516/12.
 DR N-PSDB; ABL88108.
 XX
 XX One hundred and eighty seven nucleic acids encoding PRO polypeptides, PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial PT infarction), endothelial or angiogenic disorders in a mammal -
 XX PS
 XX Claim 11; Fig 74; 565pp; English.
 XX
 XX ABL88072 to ABL88258 encode the PRO proteins given in ABB84817 to
 CC ABB85003. The PRO proteins and polynucleotides have cardiant, cytostatic,
 CC antiangiogenic, hypotensive, vulnetary and antiarteriosclerotic
 CC activities, and can be used in gene therapy. The PRO polynucleotides,
 CC proteins, agonists and antagonists are useful for treating or diagnosing
 CC a cardiovascular, endothelial or angiogenic disorder in a mammal,
 CC e.g. cardiac hypertrophy, trauma, cancer, age-related macular
 CC degeneration, atherosclerosis, hypertension, arterial restenosis,
 CC rheumatoid arthritis, angina, myocardial infarctions, thrombophlebitis,
 CC lymphangitis, tumour angiogenesis (such as breast carcinoma and liver
 CC carcinoma) and wound healing. The PRO polynucleotides have applications
 CC in molecular biology, including use as hybridisation probes, and in
 CC chromosome and gene mapping. ABL88259 to ABL88267 represent primers and
 CC probes used in the exemplification of the present invention.
 XX
 XX Sequence 867 AA;

Query Match 99.3%; Score 4719.5; DB 23; Length 867;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 866; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
 1 MGPPSLVCLLSATVFSLLGGSAFLSHRLKGRFORDRNRIRPNILVLTDQDVELGS 60
 1 MGPPSLVCLLSATVFSLLGGSAFLSHRLKGRFORDRNRIRPNILVLTDQDVELGS 60
 61 MQVNNKTRRIMEQGTHFINAFVTTMCCPSRSSILTGKVVHNHTYTNENCSSPSWQA 120
 61 MQVNNKTRRIMEQGAHFINAFVTTMCCPSRSSILTGKVVHNHTYTNENCSSPSWQA 120
 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYGSYVPPGKMWGLLKNSRFNYTLCRNGYK 180
 121 QHESRTFAVYLNSTGYRTAFPGKYLNEYGSYVPPGKMWGLLKNSRFNYTLCRNGYK 180
 181 EKHGSYSKDYLTDLTINDSVSPRTSKOMYPHREPLVLIWISHAHPGDSAPQYSRLFP 240
 181 EKHGSYSKDYLTDLTINDSVSPRTSKOMYPHREPLVLIWISHAHPGDSAPQYSRLFP 240
 241 NASQHIITPSYNYAPNDPKHIMRYTGMKPIHMEFTNMLQKRLTLMVDSDMETIYNN 300
 241 NASQHIITPSYNYAPNDPKHIMRYTGMKPIHMEFTNMLQKRLTLMVDSDMETIYNN 300
 301 LVEGTGELDNTYIVVTADHGVHIGQGLVKGSKMPEYFDIRVPYVPGPNVEAGCLNPHIV 360
 301 LVEGTGELDNTYIVVTADHGVHIGQGLVKGSKMPEYFDIRVPYVPGPNVEAGCLNPHIV 360

QY 361 LNIIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLL 420
 DB 361 LNIIDLAPTILDIAGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLL 420
 QY 421 HKRDNDKVDQAQENFELPKYQYVNDLCPQAEYQTACEQLGQKWCQVEDATGKLHKCKGP 480
 DB 421 HKRDNDKVDQAQENFELPKYQYVNDLCPQAEYQTACEQLGQKWCQVEDATGKLHKCKGP 480
 QY 481 MELGGSRALSNIPLPKYVGGSEACTCDSGDYKLSIAGRRKLLFKKKYKASVYRSRSIRSV 540
 DB 481 MELGGSRALSNIPLPKYVGGSEACTCDSGDYKLSIAGRRKLLFKKKYKASVYRSRSIRSV 540
 QY 541 AIEVDGRVYHVLGDAAPRNLTKEHWPFCAPEDQDDXGDPFSGTGGGLPDYSAANPIKVT 600
 DB 541 AIEVDGRVYHVLGDAAPRNLTKEHWPFCAPEDQDDXGDPFSGTGGGLPDYSAANPIKVT 600
 QY 601 HRCYILLENDTVOCDDLKYSLQAWKDKHLHDHETETLQNKIKNLREVRGHLKKRPEEC 660
 DB 601 HRCYILLENDTVOCDDLKYSLQAWKDKHLHDHETETLQNKIKNLREVRGHLKKRPEEC 660
 QY 661 DCHKISYHTQHKGRGLKRGSSSLHPFRKGLQEKDKVLLREQKRRKKLKLRLQNNDTC 720
 DB 661 DCHKISYHTQHKGRGLKRGSSSLHPFRKGLQEKDKVLLREQKRRKKLKLRLQNNDTC 720
 QY 721 SMPGLTCFTHDQHQWQTAPFWTLGPFCACTSANNNTYWCMETINETHNPLCFCEATGFE 780
 DB 721 SMPGLTCFTHDQHQWQTAPFWTLGPFCACTSANNNTYWCMETINETHNPLCFCEATGFE 780
 QY 781 YFDLNTDTPQLANNAVNTLDRDVLNQLHVQLMELRSCGKYKQCNPRTRNMDLGLKDGGSYE 840
 DB 781 YFDLNTDTPQLANNAVNTLDRDVLNQLHVQLMELRSCGKYKQCNPRTRNMDLGLKDGGSYE 840
 QY 841 QYRQFORRWKPEKRPSSKSLGQLWEGWEG 870
 DB 838 QYRQFORRWKPEKRPSSKSLGQLWEGWEG 867

Search completed: October 15, 2003, 13:07:55
 Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2003, 13:02:45 ; Search time 54 Seconds
(without alignments)

4157.514 Million cell updates/sec

Title: US-09-970-287-2

Perfect score: 4751
Sequence: 1 MGPPSLVLCLSATVFSLLG.....PEMKRPSKSLQLWEGWG 870

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	4751	100.0	885	Q9ULH3	Q9ulh3 homo sapien
2	4746	99.9	870	Q8IWU5	Q8iwu5 homo sapien
3	4469.5	94.1	875	Q8CFG0	Q8cfg0 mus musculu
4	4466.5	94.0	875	Q8BZ51	Q8bz51 mus musculu
5	4466.5	94.0	908	Q8BM68	Q8bm68 mus musculu
6	4462.5	93.9	875	Q8BU24	Q8bu24 mus musculu
7	4458.5	93.8	875	Q8C169	Q8cl69 mus musculu
8	3175.5	66.8	871	Q8IWU6	Q8iwu6 homo sapien
9	3142	66.1	870	Q8VI60	Q8vi60 rattus norv
10	3131.5	65.9	870	Q8K007	Q8k007 mus musculu
11	3128.5	65.8	870	Q8CLD3	Q8cld3 mus musculu
12	3117.5	65.6	867	Q9OX86	Q9ox86 coturnix co
13	3094.5	65.1	862	Q8BLJ0	Q8blj0 mus musculu
14	3059.5	64.4	604	Q8BX28	Q8bx28 mus musculu
15	3059	64.4	818	Q9UPS5	Q9ups5 homo sapien
16	2827	59.5	525	Q96SG2	Q96sg2 homo sapien

17	1914	40.3	344	4	Q9H1G9	Q9h1g9 homo sapien
18	1714.5	36.1	1114	5	Q9VEX0	Q9vex0 drosophila
19	1681	35.4	540	4	Q8NCA2	Q8nca2 homo sapien
20	1589	33.4	709	5	Q2L376	Q2l376 caenorhabdi
21	1471.5	31.0	1006	5	Q8MRG1	Q8mrg1 drosophila
22	1420	29.9	983	5	Q9N1H6	Q9n1h6 drosophila
23	1312	19.2	547	13	Q8AVH8	Q8avh8 xenopus lae
24	869.5	18.3	544	11	Q8BJJ7	Q8bjj7 mus musculu
25	868.5	18.3	544	11	Q8BK91	Q8bk91 mus musculu
26	867.5	18.3	544	11	Q8BFR4	Q8bfr4 mus musculu
27	855	18.0	149	4	Q9UTR3	Q9utr3 homo sapien
28	831	17.5	149	11	Q3D8E2	Q3d8e2 mus musculu
29	812	17.1	427	6	Q8MJB8	Q8mjb8 equus cabal
30	702	14.8	124	4	Q9H1H0	Q9h1h0 homo sapien
31	679	14.3	492	5	Q917F8	Q917f8 drosophila
32	674	14.2	492	5	Q9SR73	Q9sr73 drosophila
33	453.5	9.5	639	3	O43113	O43113 neurospora
34	430.5	9.1	649	10	O9ATG5	O9atg5 chlamydomon
35	364	7.7	517	2	Q9L5W0	Q9l5w0 prevotella
36	323	6.8	511	3	Q9P8C2	Q9p8c2 collettortic
37	301.5	6.3	497	16	Q8XC09	Q8xc09 escherichia
38	298.5	6.3	497	16	Q8PBM9	Q8pbm9 escherichia
39	291.5	6.1	169	5	Q8MVL7	Q8mvl7 boltenia vi
40	290	6.1	535	16	Q8ZHP8	Q8zhp8 versinia pe
41	279.5	5.9	491	11	Q8BL50	Q8bl50 mus musculu
42	279.5	5.9	556	11	Q9D2L1	Q9d2l1 mus musculu
43	279.5	5.9	556	11	Q8BUA1	Q8bua1 mus musculu
44	279.5	5.9	563	11	Q9CYZ0	Q9cyz0 mus musculu
45	271	5.7	562	5	Q8SZ72	Q8sz72 drosophila

ALIGNMENTS

RESULT 1

Q9ULH3 PRELIMINARY; PRT; 885 AA.
 AC Q9ULH3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein KIAA1247 (Fragment).
 GN KIAA1247.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20039619; PubMed=10574462;
 RA Nagase T., Ishikawa K., Kikuno R., Hirosewa M., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XV.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:337-345(1999).
 DR EMBL; AB033073; BAA8561.2; .
 DR InterPro; IPR002129; Pyridoxal dec.
 DR Pfam; PF00884; Sulfatase_1
 DR PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
 DR PROSITE; PS00523; SULFATASE_1; 1.
 FT NON TER 1
 SQ SEQUENCE 885 AA; 102234 MW; CCC1D4A8297AEE16 CRC64;

Query Match 100.0%; Score 4751; DB 4; Length 885;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGPPSLVLCLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRPNILVLTTDDQDVELGS 60

Db 16 MGPPSLVLCLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRPNILVLTTDDQDVELGS 75


```
Qy 61 MQVWNTTRIMEGGTHFINAFVTTIMCCPSRSSILTKGYVHNHTYTNNECSSPSWQA 120
Db 76 MQVWNTTRIMEGGTHFINAFVTTIMCCPSRSSILTKGYVHNHTYTNNECSSPSWQA 135
Qy 121 QHESRTFAVYLNSTGVRTAFPGKYLYNEYSYVPPGKSWGLLKNRPFYNTLCRNGVK 180
Db 136 QHESRTFAVYLNSTGVRTAFPGKYLYNEYSYVPPGKSWGLLKNRPFYNTLCRNGVK 195
Qy 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAHPGSDSAPQYSRLFP 240
Db 196 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAHPGSDSAPQYSRLFP 255
Qy 241 NASQHIPTSYNAPNDKHWIMRYTGMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300
Db 256 NASQHIPTSYNAPNDKHWIMRYTGMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 315
Qy 301 LVETGELDNTIYVYTADHGHIQFGLVKGSMPYBFDIRVPYVPGNVEAGCLNPHIV 360
Db 316 LVETGELDNTIYVYTADHGHIQFGLVKGSMPYBFDIRVPYVPGNVEAGCLNPHIV 375
Qy 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKQKVRWDSFLVERGKLL 420
Db 376 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKQKVRWDSFLVERGKLL 435
Qy 421 HKRDNDKVDQAEENFLPKYQVRKDLQRAEYQTAQCEQLGQKQVCVEDATGKLLHKCKGP 480
Db 436 HKRDNDKVDQAEENFLPKYQVRKDLQRAEYQTAQCEQLGQKQVCVEDATGKLLHKCKGP 495
Qy 481 MRLGGSRAALSNLVPKYVGGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540
Db 496 MRLGGSRAALSNLVPKYVGGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 555
Qy 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Db 556 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 615
Qy 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLRVGRHLKKKRPPEC 660
Db 616 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLRVGRHLKKKRPPEC 675
Qy 661 DCHKISYHTQHKGRLLKXRGSSLHPFRKGLQEKDKVWLLREQKFKKLLKRLKRLQNNDT 720
Db 676 DCHKISYHTQHKGRLLKXRGSSLHPFRKGLQEKDKVWLLREQKFKKLLKRLKRLQNNDT 735
Qy 721 SMPGLTCFTHDNQHWOTAPFTWTLGPFCACTSANNTYWCMTINETHNLFCEPATGFLE 780
Db 736 SMPGLTCFTHDNQHWOTAPFTWTLGPFCACTSANNTYWCMTINETHNLFCEPATGFLE 795
Qy 781 YFDLNTDPYOLMNAVNTLDRDVLNQLHVQWELRSCGYKQCNPRTRNNMDLGLKDGGSYE 840
Db 796 YFDLNTDPYOLMNAVNTLDRDVLNQLHVQWELRSCGYKQCNPRTRNNMDLGLKDGGSYE 855
Qy 841 QYRQFORRWKPEMKRPSKSLGOLWEGWEG 870
Db 856 QYRQFORRWKPEMKRPSKSLGOLWEGWEG 885
```

RESULT 2

```
Q8IWUS PRELIMINARY; PRT: 870 AA.
AC Q8IWUS,
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Extracellular sulfatase SULF-2.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
```

```
RX PubMed=12368295;
RA Morimoto-Tomita M., Uchimura K., Werb Z., Hemmerich S., Rosen S.D.;
RT "Cloning and Characterization of Two Extracellular Heparin-degrading
RT Endosulfatases in Mice and Humans.";
EL J. Biol. Chem. 277:49175-49185 (2002).
DR EMBL; AY101176; AA076861.1; -. 74B1069CE2774D73 CRC64;
SQ SEQUENCE 870 AA; 100434 MW; 74B1069CE2774D73 CRC64;

Query Match 99.9%; Score 4746; DB 4; Length 870;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 869; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPPSLVLCLLSATVFSLLGSSAFISHRLKGRFORDRRNIRPNIIILVTDQDVELGS 60
Db 1 MGPPSLVLCLLSATVFSLLGSSAFISHRLKGRFORDRRNIRPNIIILVTDQDVELGS 60
Qy 61 MQVWNTTRIMEGGTHFINAFVTTIMCCPSRSSILTKGYVHNHTYTNNECSSPSWQA 120
Db 61 MQVWNTTRIMEGGTHFINAFVTTIMCCPSRSSILTKGYVHNHTYTNNECSSPSWQA 120
Qy 121 QHESRTFAVYLNSTGVRTAFPGKYLYNEYSYVPPGKSWGLLKNRPFYNTLCRNGVK 180
Db 121 QHESRTFAVYLNSTGVRTAFPGKYLYNEYSYVPPGKSWGLLKNRPFYNTLCRNGVK 180
Qy 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAHPGSDSAPQYSRLFP 240
Db 181 EKHGSDYSKDYLTDLITNDSVSFFRTSKMYPHRPVLMVISHAAHPGSDSAPQYSRLFP 240
Qy 241 NASQHIPTSYNAPNDKHWIMRYTGMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300
Db 241 NASQHIPTSYNAPNDKHWIMRYTGMKPIHMEFTNMLQRLQTLMSVDDSMETIYNN 300
Qy 301 LVETGELDNTIYVYTADHGHIQFGLVKGSMPYBFDIRVPYVPGNVEAGCLNPHIV 360
Db 301 LVETGELDNTIYVYTADHGHIQFGLVKGSMPYBFDIRVPYVPGNVEAGCLNPHIV 360
Qy 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKQKVRWDSFLVERGKLL 420
Db 361 LNTDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKQKVRWDSFLVERGKLL 420
Qy 421 HKRDNDKVDQAEENFLPKYQVRKDLQRAEYQTAQCEQLGQKQVCVEDATGKLLHKCKGP 480
Db 421 HKRDNDKVDQAEENFLPKYQVRKDLQRAEYQTAQCEQLGQKQVCVEDATGKLLHKCKGP 480
Qy 481 MRLGGSRAALSNLVPKYVGGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540
Db 481 MRLGGSRAALSNLVPKYVGGSEACTCDSGDYKLSLAGRRKKLFKKYKASYVRSRSIRSV 540
Qy 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Db 541 AIEVDGRVYHVLGDAAPRNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAANPIKVT 600
Qy 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLRVGRHLKKKRPPEC 660
Db 601 HRCYILENDTVQCDLDLYKSLQAWKDKHLHIDHEIETLQNKIKNLRVGRHLKKKRPPEC 660
Qy 661 DCHKISYHTQHKGRLLKXRGSSLHPFRKGLQEKDKVWLLREQKFKKLLKRLKRLQNNDT 720
Db 661 DCHKISYHTQHKGRLLKXRGSSLHPFRKGLQEKDKVWLLREQKFKKLLKRLKRLQNNDT 720
Qy 721 SMPGLTCFTHDNQHWOTAPFTWTLGPFCACTSANNTYWCMTINETHNLFCEPATGFLE 780
Db 721 SMPGLTCFTHDNQHWOTAPFTWTLGPFCACTSANNTYWCMTINETHNLFCEPATGFLE 780
Qy 781 YFDLNTDPYOLMNAVNTLDRDVLNQLHVQWELRSCGYKQCNPRTRNNMDLGLKDGGSYE 840
Db 781 YFDLNTDPYOLMNAVNTLDRDVLNQLHVQWELRSCGYKQCNPRTRNNMDLGLKDGGSYE 840
Qy 841 QYRQFORRWKPEMKRPSKSLGOLWEGWEG 870
Db 841 QYRQFORRWKPEMKRPSKSLGOLWEGWEG 870
```

```

RESULT 3
Q8CFG0
ID Q8CFG0 PRELIMINARY; PRT; 875 AA.
AC Q8CFG0;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Extracellular sulfatase SULF-2.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gross mammary tumor;
RX PubMed=12368295;
RA Morimoto-Tomita M., Uchimura K., Werb Z., Hemmerich S., Rosen S.D.,
RT "Cloning and Characterization of Two Extracellular Heparin-degrading
RT Endosulfatases in Mice and Humans";
RL J. Biol. Chem. 277:49175-49185 (2002).
DR EMBL: AY101117; AAM76862.1; -.
SQ SEQUENCE 875 AA; 100469 MW; 148406BCCC7D76B9 CRC64;

Query Match 94.1%; Score 4469.5; DB 11; Length 875;
Best Local Similarity 93.1%; Pred. No. 0;
Matches 821; Conservative 18; Mismatches 30; Indels 7; Gaps 3;

QY 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNPNIILVLTDDQDVELGS 60
Db 1 MAPPGPLWLLSTALLSLLAGSSAFSLSHRLKGRFQDRNRIRNPNIILVLTDDQDVELGS 60
QY 61 MQVMKTRIMEGGTHFINAFVTPPMCCPSRSSILTKGYVHNHTYTNNECSPSQWA 120
Db 61 MQVMKTRIMEGGTHFINAFVTPPMCCPSRSSILTKGYVHNHTYTNNECSPSQWA 120
QY 121 QHESRTFAVILNSTGYRTAFGKYLNEYNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 180
Db 121 QHESRTFAVILNSTGYRTAFGKYLNEYNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 180
QY 181 EXHGSYSKDYLTDLITNDVSFFRTSKMYPHPRVPLVMSHAAPHGDSAPQYSRLFP 240
Db 181 EXHGSYSKDYLTDLITNDVSFFRTSKMYPHPRVPLVMSHAAPHGDSAPQYSRLFP 240
QY 241 NASQHTPSYVAPNDKHWIMRYTGMKPIHMEFTNMLQRKLTQLMSVDDSMETIYNM 300
Db 241 NASQHTPSYVAPNDKHWIMRYTGMKPIHMEFTNMLQRKLTQLMSVDDSMETIYNM 300
QY 301 LVETGELDNITYIYVYADHGYHIGFGLVKGSMPEYFDIRVPFYVGPVNEAGSLNPHIV 360
Db 301 LVETGELDNITYIYVYADHGYHIGFGLVKGSMPEYFDIRVPFYVGPVNEAGSLNPHIV 360
QY 361 LVETGELDNITYIYVYADHGYHIGFGLVKGSMPEYFDIRVPFYVGPVNEAGSLNPHIV 420
Db 361 LVETGELDNITYIYVYADHGYHIGFGLVKGSMPEYFDIRVPFYVGPVNEAGSLNPHIV 420
QY 421 HKRDNKDVAQENFLPKYQVVDLCORAEYQTACEQLGKQKQCVEDATGKLKHKCKGP 480
Db 421 HKRDNKDVAQENFLPKYQVVDLCORAEYQTACEQLGKQKQCVEDATGKLKHKCKGP 480
QY 481 MRL- --GGSRALSNLPYKYGGSSEACTCDS- --GDYKLSLAGRRKLPKKYKASVYRS 534
Db 481 MRL- --GGSRALSNLPYKYGGSSEACTCDS- --GDYKLSLAGRRKLPKKYKASVYRS 534
QY 535 RSIRSAIEVDGRVYHVLGDAQAQPNRLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAA 594
Db 535 RSIRSAIEVDGRVYHVLGDAQAQPNRLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAA 594
QY 540 RSIRSAIEVDGEIYHVGLDTPQPNLSKPHWPGAPEDQDDKGGDFSGTGGLPDYSAP 599
Db 540 RSIRSAIEVDGEIYHVGLDTPQPNLSKPHWPGAPEDQDDKGGDFSGTGGLPDYSAP 599
QY 595 NP1KVTNRCVILENDTVQCDLDLYKSLQAKDKHLHIHIEITLQNKIKNLEVRGHLKK 654
Db 595 NP1KVTNRCVILENDTVQCDLDLYKSLQAKDKHLHIHIEITLQNKIKNLEVRGHLKK 654
QY 600 NP1KVTNRCVILENDTVQCDLDLYKSLQAKDKHLHIHIEITLQNKIKNLEVRGHLKK 659
Db 600 NP1KVTNRCVILENDTVQCDLDLYKSLQAKDKHLHIHIEITLQNKIKNLEVRGHLKK 659
QY 655 KRPEECDCHKISYTHQKGRKLRKHKSSLHPRFKGLQEKDKVLLRQEKKKLKLKRL 714
Db 655 KRPEECDCHKISYTHQKGRKLRKHKSSLHPRFKGLQEKDKVLLRQEKKKLKLKRL 714

RESULT 4
Q8BZ51
ID Q8BZ51 PRELIMINARY; PRT; 875 AA.
AC Q8BZ51;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Weakly similar to N-acetylglucosamine-6-sulfatase.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL: AK036685; BAC229534.1; -.
SQ SEQUENCE 875 AA; 100497 MW; A04AB401029FCCDC CRC64;

Query Match 94.0%; Score 4466.5; DB 11; Length 875;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 820; Conservative 19; Mismatches 30; Indels 7; Gaps 3;

QY 1 MGPPSLVCLLSATVFSLLGSSAFSLSHRLKGRFQDRNRIRNPNIILVLTDDQDVELGS 60
Db 1 MAPPGPLWLLSTALLSLLAGSSAFSLSHRLKGRFQDRNRIRNPNIILVLTDDQDVELGS 60
QY 61 MQVMKTRIMEGGTHFINAFVTPPMCCPSRSSILTKGYVHNHTYTNNECSPSQWA 120
Db 61 MQVMKTRIMEGGTHFINAFVTPPMCCPSRSSILTKGYVHNHTYTNNECSPSQWA 120
QY 121 QHESRTFAVILNSTGYRTAFGKYLNEYNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 180
Db 121 QHESRTFAVILNSTGYRTAFGKYLNEYNGSVYPPGKWEVGLLKNRFFNYTLCRNGVK 180
QY 181 EXHGSYSKDYLTDLITNDVSFFRTSKMYPHPRVPLVMSHAAPHGDSAPQYSRLFP 240
Db 181 EXHGSYSKDYLTDLITNDVSFFRTSKMYPHPRVPLVMSHAAPHGDSAPQYSRLFP 240
QY 241 NASQHTPSYVAPNDKHWIMRYTGMKPIHMEFTNMLQRKLTQLMSVDDSMETIYNM 300
Db 241 NASQHTPSYVAPNDKHWIMRYTGMKPIHMEFTNMLQRKLTQLMSVDDSMETIYNM 300
QY 301 LVETGELDNITYIYVYADHGYHIGFGLVKGSMPEYFDIRVPFYVGPVNEAGSLNPHIV 360
Db 301 LVETGELDNITYIYVYADHGYHIGFGLVKGSMPEYFDIRVPFYVGPVNEAGSLNPHIV 360
QY 361 LVETGELDNITYIYVYADHGYHIGFGLVKGSMPEYFDIRVPFYVGPVNEAGSLNPHIV 420
Db 361 LVETGELDNITYIYVYADHGYHIGFGLVKGSMPEYFDIRVPFYVGPVNEAGSLNPHIV 420
QY 421 HKRDNKDVAQENFLPKYQVVDLCORAEYQTACEQLGKQKQCVEDATGKLKHKCKGP 480
Db 421 HKRDNKDVAQENFLPKYQVVDLCORAEYQTACEQLGKQKQCVEDATGKLKHKCKGP 480

```

```

Db 421 HKREGDKVNAQENFLPKYQRYKDLQRAEYQTAQCEQLGQKQVCVEDASGTLKLHKCKGP 480
QY 481 MRL---GGSRALSNLVPKYGGSEACTCDS---GDYKLSLAGRRKKLFKKYKASYVRS 534
Db 481 MRFGGGSSRALSNLVPKYDQSSSEACSGGGDYKLGAGRR-KLFKKYKTSYARN 539
QY 535 RSIRSAIEVDGRVYHVGLGDAQAQPNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAA 594
Db 540 RSIRSAIEVDGEIYHVGLDTVPQPNLSKPHWPGAPEDQDDKGGDFSGTGGLPDYSAP 599
QY 595 NPIKVTHRCYILENDTVQCDLDLYKSLQAWKHJLHDHEIETLQNKIKNLREVGHLLK 654
Db 600 NPIKVTHRCYILENDTVQCDLDLYKSLQAWKHJLHDHEIETLQNKIKNLREVGHLLK 659
QY 655 KRPECDCHKISYHTQHKGRSLHGRSSLHPPFRKGLQEKOKVLLRQKRRKKLKLKLL 714
Db 660 KRPECDCHKISYHSQHKGRSLHGRSSLHPPFRKGLQEKOKVLLRQKRRKKLKLKLL 719
QY 715 QNNDTCSMPGLTCFTHDNQHWQTAFTLGPFCACTSANNNTYWCMTTINETHNPLFCBP 774
Db 720 QNNDTCSMPGLTCFTHDNQHWQTAFTLGPFCACTSANNNTYWCMTTINETHNPLFCBP 779
QY 775 ATGFLEYFDLNTDYPQLMNAVNTLDRVLNQLHVMELRSCKYKOCNPRTRNMDLGLK 834
Db 780 ATGFLEYFDLNTDYPQLMNAVNTLDRVLNQLHVMELRSCKYKOCNPRTRNMDLGLR 839
QY 835 DGSSEYEQYQFORRKWPKMKRPSKSLGQLWEGWEG 870
Db 840 DGSSEYEQYQFORRKWPKMKRPSKSLGQLWEGWEG 875

RESULT 5
QB8M68
ID QB8M68 PRELIMINARY; PRT; 908 AA.
AC QB8M68;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Weakly similar to N-acetylglucosamine-6-sulfatase.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
DR EMBL; AK034712; BAC28804.1; -.
SQ SEQUENCE 908 AA; 104037 MW; 1B460655AEC3734D CRC64;

Query Match 94.0%; Score 4466.5; DB 11; Length 908;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 820; Conservative 19; Mismatches 30; Indels 7; Gaps 3;

QY 1 MGPPSLVLLSATVFLSGSSAFSLHRLKGRFQDRDRNIRPNILVLTDQDVELGS 60
Db 34 MAPGCLPLWLJSTALLSLLAGSFAFLSHPRLKGRFQDRDRNIRPNILVLTDQDVELGS 93
QY 61 MOVNKTTRIMEQGTTHFINAVFTTPMCCPSRSSILTKGVYVHNHTYNNENCSSPSWA 120
Db 94 MOVNKTTRIMEQGAHFINAVFTTPMCCPSRSSILTKGVYVHNHTYNNENCSSPSWA 153
QY 121 CHESTFAVLNLTGTYRTAFFGKLYNEYSYVPPGKWKWGLLQNSRFNYTLCRNGVK 180
Db 154 CHESTFAVLNLTGTYRTAFFGKLYNEYSYVPPGKWKWGLLQNSRFNYTLCRNGVK 213
QY 181 EKHGSDYSDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHGPDPSAPQYSLRFP 240

```

```

Db 214 EKHGSDYSDYLTDLITNDVSFFRTSKMYPHRPVLMVISHAAPHGPDPSAPQYSLRFP 273
QY 241 NASOHLTPSYNAPNPDGKIMRYTGPMPKPHMETFNNLQKRLOTLSVDDSMETIYNM 300
Db 274 NASOHLTPSYNAPNPDGKIMRYTGPMPKPHMETFNNLQKRLOTLSVDDSMETIYDM 333
QY 301 LVETGELDNITYIVYTADHGYHIGQFGLVKGSMPEYFDIRVPFYVYRGNVAGCLNPHIV 360
Db 334 LVETGELDNITYILYTADHGYHIGQFGLVKGSMPEYFDIRVPFYVYRGNVAGCLNPHIV 393
QY 361 LNIDLAPTLDIAGLDIPADMGGKSLILKLDTERPVNPFHLKKORVWDSFLVERGKLL 420
Db 394 LNIDLAPTLDIAGLDIPADMGGKSLILKLDTERPVNPFHLKKORVWDSFLVERGKLL 453
QY 421 HKRDNDKVDQAQENFLPKYQRYKDLQRAEYQTAQCEQLGQKQVCVEDATGKLKLHKCKGP 480
Db 454 HKREGDKVNAQENFLPKYQRYKDLQRAEYQTAQCEQLGQKQVCVEDASGTLKLHKCKGP 513
QY 481 MRL---GGSRALSNLVPKYGGSEACTCDS---GDYKLSLAGRRKKLFKKYKASYVRS 534
Db 514 MRFGGGSSRALSNLVPKYDQSSSEACSGGGDYKLGAGRR-KLFKKYKTSYARN 572
QY 535 RSIRSAIEVDGRVYHVGLGDAQAQPNLTKRHWPGAPEDQDDKGGDFSGTGGLPDYSAA 594
Db 573 RSIRSAIEVDGEIYHVGLDTVPQPNLSKPHWPGAPEDQDDKGGDFSGTGGLPDYSAP 632
QY 595 NPIKVTHRCYILENDTVQCDLDLYKSLQAWKHJLHDHEIETLQNKIKNLREVGHLLK 654
Db 633 NPIKVTHRCYILENDTVQCDLDLYKSLQAWKHJLHDHEIETLQNKIKNLREVGHLLK 692
QY 655 KRPECDCHKISYHTQHKGRSLHGRSSLHPPFRKGLQEKOKVLLRQKRRKKLKLKLL 714
Db 693 KRPECDCHKISYHSQHKGRSLHGRSSLHPPFRKGLQEKOKVLLRQKRRKKLKLKLL 752
QY 715 QNNDTCSMPGLTCFTHDNQHWQTAFTLGPFCACTSANNNTYWCMTTINETHNPLFCBP 774
Db 753 QNNDTCSMPGLTCFTHDNQHWQTAFTLGPFCACTSANNNTYWCMTTINETHNPLFCBP 812
QY 775 ATGFLEYFDLNTDYPQLMNAVNTLDRVLNQLHVMELRSCKYKOCNPRTRNMDLGLK 834
Db 813 ATGFLEYFDLNTDYPQLMNAVNTLDRVLNQLHVMELRSCKYKOCNPRTRNMDLGLR 872
QY 835 DGSSEYEQYQFORRKWPKMKRPSKSLGQLWEGWEG 870
Db 873 DGSSEYEQYQFORRKWPKMKRPSKSLGQLWEGWEG 908

RESULT 6
QB8U24
ID QB8U24 PRELIMINARY; PRT; 875 AA.
AC QB8U24;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Weakly similar to N-acetylglucosamine-6-sulfatase.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
DR EMBL; AK081843; BAC38279.1; -.
SQ SEQUENCE 875 AA; 100531 MW; 4D05AC0923B026F4 CRC64;

Query Match 93.9%; Score 4462.5; DB 11; Length 875;
Best Local Similarity 93.5%; Pred. No. 0;

```



```
QY 835 DGSYQYQFQRRKWPMPKRPSSKSLGQWEGWEG 870
Db 840 DGSYQYQFQRRKWPMPKRPSSKSLGQWEGWEG 875

RESULT 8
Q81WU6 PRELIMINARY; PRT; 871 AA.
AC Q81WU6;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DE Extracellular sulfatase SULF-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP PubMed=12368295;
RA Morimoto-Tonita M., Uchimura K., Werb Z., Hemmerich S., Rosen S.D.;
RT "Cloning and Characterization of Two Extracellular Heparin-degrading
RT Endosulfatases in Mice and Humans.";
RL J. Biol. Chem. 277:49175-49185 (2002).
DR EMBL; AY101175; AAW76860.1; -.
SQ SEQUENCE 871 AA; 101026 MW; 9A90ADB280304364 CRC64;

Query Match 66.1%; Score 3175.5; DB 4; Length 871;
Best Local Similarity 64.3%; Pred. No. 1.5e-244;
Matches 578; Conservative 120; Mismatches 142; Indels 51; Gaps 12;

QY 5 SLVCLLSATVFSLLGSSAFSLSHRLKGRFDRRNPRIILVLTDDQDVELGSLQVM 64
Db 7 ALLAVLGT---ELLGSLCSTVSPRFRGRIQOERNIRPNIIILVLTDDQDVELGSLQVM 63
QY 65 NKTRRIMEOGTHFINAFVTTMCCPSRSSILTGVVHNHTYTNNECSSPSWQAQHES 124
Db 64 NKTRKMEHGGAFTINAFVTTMCCPSRSSILTGVVHNHTYTNNECSSPSWQAQHEP 123
QY 125 RTFAVYLNSTGYRTAFPGKYLNEYNGSYPPGKWEVGLLKNRSFYNYTLCRNGVKEKHG 184
Db 124 RTFAVYLNNTGYRTAFPGKYLNEYNGSYPPGKWEVGLLKNRSFYNYTVCRCNGVKEKHG 183
QY 185 SDYKDYLTDLTINDSVSPRFTSKMYPHRPVLMVISHAAPHGPDSPAPQSKLPNNAQ 244
Db 184 FDYKDYFTDLITNESINTFYKSKMYPHRPVLMVISHAAPHGPDSPAPQSKLPNNAQ 243
QY 245 HITPSYNYAPNDKHWIMRYTGPMKPIHMEFTNMLQRLQTLMSYDSSMETIYNMLVET 304
Db 244 HITPSYNYAPNDKHWIMRYTGPMKPIHMEFTNMLQRLQTLMSYDSSVERLYNMLVET 303
QY 305 GELDNTYIVYTAHGYHIGQFGLVKGKSMYPEDIRVPYVGRPNVEAGCLNPHIVLNI 364
Db 304 GELNTYIYTAHGYHIGQFGLVKGKSMYPEDIRVPYVGRPNVEAGCLNPHIVLNI 363
QY 365 LAPTILDIAGLDIPADMDGSKILKLDTERPNVRPHLKKMRVWRDSFVVERGKLLHKRD 424
Db 364 LAPTILDIAGLDTPPDVGGKSLKLDPEKPGNRFRNKKAKIWRDTPFLVERGKELRKE 423
QY 425 NDYKDAQEEFLPKYQVXDLCORAYOTACEOLGQKWCQVEDATGKLLKHKCKGPMRLG 484
Db 424 ESSKNIQOSSLNPKYERVKELCOQARYOTACEOGKQWCQIEDTSGLKIRHKCKGPDLL 483
QY 485 GSR-ALSNLPKYVGQSEACTCDSDGYKLSLAGRR-KKLFFK---KYKASYRYSRI 537
Db 484 TVRQSTNLLVARGFHDKDECCRESGYRASRSQRKSQRFLENQTPKPKRPFVHTROT 543
QY 538 RSVIAEVDGRVYHVLGSD-----AAQPRNLTKRWFC--APEDQDDKGG-----DFS 583
Db 544 RSLSEVEGEIYDINLEEBELQVLPRIAKRGHKGPRDLQASSGNGRMLADSS 603
QY 584 GTGGLPDYSAANPKVTHRCYIILENDTVQCDLLDYKSLQAWKHLDIHEIETIQNKIK 643
```

```
Db 604 NAVGPP-----TTVRTHKCFILPNDSTHCERELQSAKAWKHKAYIDKEIALQDKIK 658
QY 644 NLREVRGHLKKRPBECCHKISYHTQHKRLKHR--GSSLHPPFRKGLQEKD-KYMLLRE 700
Db 659 NLREVRGHLKKRPBECSCSKSYNKGKGVKQEKLKSHLHPFKAAQAEVDSKLQLPKE 718
QY 701 -OKRKKKRLKRLQNNDTCSMPGLTCTFDHNOHWQTAPFWTLGPPCACTSANNTYWC 759
Db 719 NNRERRKKEKRRQKGBECSLPGLTCTFDHNNHWQTAPFWNLGSCFCACTSSNNNTYWC 778
QY 760 MRTINETHFLCEPATGPLEYFDLNTDPYQLMAVNTLDRDVLNQLHVLQMLRSCYKY 819
Db 779 LRTVNEHNLFCFATGPLEYFDLNTDPYQLTNTVHTVERGILNQLHVLQMLRSCQY 838
QY 820 KQCNPRTRNMDLGLGKGSYEQYRQRRKWPMPKRPSSKSLGQWEGWEG 870
Db 839 KQCNPRPKNLDVGNKDGSSYDLR-----GQLWDGWEW 871

RESULT 9
Q8VI60 PRELIMINARY; PRT; 870 AA.
AC Q8VI60;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Sulfatase PP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
SEQUENCE FROM N.A.
RA Ohto T., Uchida H., Yamazaki H., Keino-Masu K., Masu M.;
RT "Identification of a novel non-lysosomal sulfatase expressed in the
RT floor plate, choroid plexus, and cartilage.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF230072; NAL71906.1; -.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
SQ SEQUENCE 870 AA; 100859 MW; 037813868C08967F CRC64;

Query Match 66.1%; Score 3142; DB 11; Length 870;
Best Local Similarity 64.3%; Pred. No. 7e-242;
Matches 572; Conservative 124; Mismatches 144; Indels 50; Gaps 11;

QY 5 SLVCLLSATVFSLLGSSAFSLSHRLKGRFDRRNPRIILVLTDDQDVELGSLQVM 64
Db 7 ALLAVLGA---ELLGPGCSTWRSQRFRGRVQOERKNIRPNIIILVLTDDQDVELGSLQVM 63
QY 65 NKTRRIMEOGTHFINAFVTTMCCPSRSSILTGVVHNHTYTNNECSSPSWQAQHES 124
Db 64 NKTRKMEHGGAFTINAFVTTMCCPSRSSILTGVVHNHTYTNNECSSPSWQAQHEP 123
QY 125 RTFAVYLNSTGYRTAFPGKYLNEYNGSYPPGKWEVGLLKNRSFYNYTLCRNGVKEKHG 184
Db 124 RTFAVYLNNTGYRTAFPGKYLNEYNGSYPPGKWEVGLLKNRSFYNYTVCRCNGVKEKHG 183
QY 185 SDYKDYLTDLTINDSVSPRFTSKMYPHRPVLMVISHAAPHGPDSPAPQSKLPNNAQ 244
Db 184 FDYKDYFTDLITNESINTFYKSKMYPHRPVLMVISHAAPHGPDSPAPQSKLPNNAQ 243
QY 245 HITPSYNYAPNDKHWIMRYTGPMKPIHMEFTNMLQRLQTLMSYDSSMETIYNMLVET 304
Db 244 HITPSYNYAPNDKHWIMRYTGPMKPIHMEFTNMLQRLQTLMSYDSSVERLYNMLVET 303
QY 305 GELDNTYIVYTAHGYHIGQFGLVKGKSMYPEDIRVPYVGRPNVEAGCLNPHIVLNI 364
Db 304 GELNTYIYTAHGYHIGQFGLVKGKSMYPEDIRVPYVGRPNVEAGCLNPHIVLNI 363
QY 365 LAPTILDIAGLDIPADMDGSKILKLDTERPNVRPHLKKMRVWRDSFVVERGKLLHKRD 424
```

```

Db 364 LAPILDIAGLDPSPVDGSKVLKLDLEKPGNRFRTNKAKIMRDTFLVERGKFLRKE 423
Qy 425 NDKVDAQEENFLPKYORVKDLCORAEYQACEQGWOCQVEDATGKLKHKCKGPMRLG 484
Db 424 ESSKNIOQSNHLPKYERVKELCOOARYQACEOPGQNMOCIEDTSGLRIHKCKGPDLL 483
Qy 485 GSRA-LSNLVPKYGGSEACTCDSGDYKLSLGRPK-KLFKK-----KYKASYVRSR1 537
Db 484 TVRQANRLYSRGLQDKDKECHRESGYRPSRQRKQFLRNOGTQPKYPRFVHTRT 543
Qy 538 RSVAEIVDGRVYHVLGD-----AAQPNLTKRHWP-----APEDDDKGGDFSG 584
Db 544 RSLSVFEGEYDINLEEBELQVLPFRSAKHDEGHQFCHQAAAGDIRMELADSN 603
Qy 585 TGLDPSAANPIKVTIRCVILENDTVQCDLDYKSLQAWKHUHIHEITLQNKIN 644
Db 604 AVGLP-----TVRVTHKCFILPNDTHCERELYQSARAWKHAYIDKEIEVLQDKIN 658
Qy 645 LREVRGHLKKRPEECCHKISYHTQHG--RLKHGSSLHPFRK-GLOEKD-KVWLARE 700
Db 659 LREVRGHLKKRPEECGQSYNKEKGVKROEKLSHLHPFKAAAEVDSKLQLFKE 718
Qy 701 QKXKGLKRLKLLKQLNNDTCSMPGLTCFTHDNHWTAPFWTLGPFCACTSANNTYWC 760
Db 719 HRRKKERKEKQKRGECESLPGLTCFTHDNHWTAPFWNLGSPCACTSSNNNTYWC 778
Qy 761 RTINETHNLFCEFPATGFLYFDLNTDPVQLMNAVNTLDRDLNQLHVLQMLRSCGYK 820
Db 779 RTVNETHNFCEFPATGFLYFDLNTDPVQLMNAVNTLDRDLNQLHVLQMLRSCGYK 838
Qy 821 QCNPRTRNDLGLKGGSYEQYQFORRWKPEKRPSSKSLGQWEGWEG 870
Db 839 QCNPRKSLDVGTEKGGNYDPR-----GQLWDGWEG 870

```

RESULT 10

```

Q8K007 ID Q8K007 PRELIMINARY; PRT; 870 AA.
AC 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Similar to sulfatase FP (Extracellular sulfatase SULF-1).
GN AW121680.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX PubMed=12368295;
RA Morimoto-Tomita M., Uchimura K., Werb Z., Hemmerich S., Rosen S.D.;
RT "Cloning and Characterization of Two Extracellular Heparin-degrading
RT Endosulfatases in Mice and Humans.";
J. Biol. Chem. 277:49175-49185 (2002).
DR EMBL; BC034547; AAH34547.1; -
DR EMBL; AY101178; AAH76863.1; -
DR MGD; MGI:2138563; AW121680.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE 1; 1.
SQ SEQUENCE 870 AA; 100922 MW; 4A9BE710D7CF4F9D CRC64;

```

Query Match 65.9%; Score 3131.5; DB 11; Length 870;
 Best Local Similarity 64.2%; Pred. No. 4.8e-241;
 Matches 571; Conservative 122; Mismatches 150; Indels 47; Gaps 10;

RESULT 11

```

Q8C1D3 ID Q8C1D3 PRELIMINARY; PRT; 870 AA.
AC 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Sulfatase FP homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```



```

Db 544 RLSVFEFEIYDINLEEEELQVLTSTRSITKRNHNDKKAETDGGAPGDGMWADGTVI 603
QY 584 GTGGLPDYSAANPIKVTTHRCYILENDTVQCDLDLYKSLQAWKHLDHIEITLQNKIK 643
Db 604 G-----QPSSVVRVTHKCFILPNDTTCERELYSARAKWKHAYIDKXIEALQDKIK 655
QY 644 NUREVRGHLKKRPEECDCCHKISYTHQHG-----RLKHGSSSLHPPFRKGLQEKD-KWVL 697
Db 656 NUREVRGHLKKRPEECDCCHKISYTHQHG-----SHLHPFKAAQEVDSKLQL 712
QY 698 LREOKRKKLKKLKLQNDTCSMPGLTCFTHDNOHWOTAPFTWLGPFCACTSANNTY 757
Db 713 FENRRKKERKKRQKKGDECSLPGLTCFTHDNNHWOTAPFWNLGSFCACTSSNNTY 772
QY 758 WCMRTINETHNLFCEFAFGLEYFDLNTDPYQLMNAVNTLDRDLVNLQHLVQLMELRSCK 817
Db 773 WCLRTVNDTHNLFCEFAFGLEYFDLNTDPYQLMNAVNTLDRDLVNLQHLVQLMELRSCK 832
QY 818 GYKQCNPRTRNDLGLKDGGSYEYQVRQRRKWPMPKRPSSKSLGOLWEGWEG 870
Db 833 GYKQCNPRPKLETGNKDGGSYDPRH-----GQLMDGWEH 867

RESULT 13
Q8BLJ0 PRELIMINARY; PRT; 862 AA.
AC Q8BLJ0;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Sulfatase Fp homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).";
DR EMBL; AK045002; BAC32179.1; --
SQ SEQUENCE 862 AA; 9964 MW; BE67B180D3D26359 CRC64;

Query Match 65.1%; Score 3094.5; DB 11; Length 862;
Best Local Similarity 65.2%; Pred. No. 4.3e-238;
Matches 563; Conservative 121; Mismatches 150; Indels 29; Gaps 9;

QY 5 SLVCLLSATVFSLLGSSAFLSHRLKGRFORDRNRIRPNILVLTDQDVELGSMQVM 64
Db 4 SLWALLAVLGTQLLGLSLCSTVRSQRFGRIOQRKNIRPNILVLTDQDVELGSLQVM 63
QY 65 NKTRIMEQGGTHFINAFVTTCPPSRSSILTKYVHNHNTYNNENCSFSPQAOHES 124
Db 64 NKTRIMEQGGATFTNAFTVTTMCCPSRSSMLTKYVHNHNTYNNENCSFSPQAOHEP 123
QY 125 RTFAVYLNSTGYRTAFEGKLYNEYSYVPPGWKEWGLKXSRFYNTLCRNGVKEKHG 184
Db 124 RTFAVYLNSTGYRTAFEGKLYNEYSYVPPGWKEWGLKXSRFYNTLCRNGVKEKHG 183
QY 185 SDYSKDYLTDLITNDSVFSRTSKQYRPHRVLWVISHAHPGSDSAPQYSRLFPNASQ 244
Db 184 FDYAKDYFTDLITNESINYFNQSKRMYPHRPIMWVISHAHPGSDSAPQYSRLFPNASQ 243
QY 245 HITPSYNAIPNPKHMYTGMKDIHMEFTNMLQRLQTLMSVDDSMETIYNMLVET 304
Db 244 HITPSYNAIPNPKHMYTGMKDIHMEFTNMLQRLQTLMSVDDSMETIYNMLVET 303
QY 305 GELDNTYIVYADHGHIQFGLVKGKSMYPEDIRVPFVVRGPNVEAGCLNPHVILNID 364

```

```

Db 304 GELDNTYIVYADHGHIQFGLVKGKSMYPEDIRVPFVVRGPNVEAGCLNPHVILNID 363
QY 365 LAPTILDIAGLDIPADMDGKSLKLLDTEREVNRFHLKCKOORVWRDSFLVERGKLLHKRD 424
Db 364 LAPTILDIAGLDSPSDVDGKSVLKLDLEKPGNFRNTKKAKIWRDTFLVERGKFLAKE 423
QY 425 NDKVDAQENFLPKYQVRVDLCQRAEYOTACEQLQKQWQCVEDATGKLKHKCKGPNRLG 484
Db 424 ESKNIOQSNHLPKYRKYKELCQQAQYOTACEQPGQNMQCIEDTSGKLRHKCKGPNRLG 483
QY 485 GSRA-LSNLVPKYGGGSEACTCSDGYKLSLAGR---RKCLPKK---KYKASVYRSRS1 537
Db 484 TVRQARNLYSRGLHDKDKCHCRSDGYRSRSRQKQORQFLRNKGTPKYKPRFVHTQT 543
QY 538 RVALTEVDGRVTHVGLGD-----AAQPNMLTGRHWPQ-----APEDODDDKGGDFSG 584
Db 544 RLSVFEFEIYDINLEEEELQVLPKRSIAKRHDGEGHGFIGHQAAAGDIRNEMLAASN 603
QY 585 TGGLPDYSAANPIKVTTHRCYILENDTVQCDLDLYKSLQAWKHLDHIEITLQNKIK 644
Db 604 AVGLP-----ATVRVTHKCFILPNDTTCERELYSARAKWKHAYIDKXIEALQDKIK 658
QY 645 LREVRGHLKKRPEECDCCHKISYTHQHG--RLKHGSSSLHPPFRK-GLQEKD-KWVLRE 700
Db 659 LREVRGHLKKRPEECDCCHKISYTHQHG--RLKHGSSSLHPPFRK-GLQEKD-KWVLRE 718
QY 701 OKRKKKRLKRLKLNNDTCSMPGLTCFTHDNOHWOTAPFTWLGPFCACTSANNTYCM 760
Db 719 HRRKKRKKRQKRGBCSLPGLTCFTHDNNHWOTAPFWNLGSFCACTSSNNTYACL 778
QY 761 RTINETHNLFCEFAFGLEYFDLNTDPYQLMNAVNTLDRDLVNLQHLVQLMELRSCKGYK 820
Db 779 RTVNETHNLFCFAFGLEYFDLNTDPYQLMNAVNTLDRDLVNLQHLVQLMELRSCKGYK 838
QY 821 QCNPRTRNDLGLKDGGSYEYQVR 843
Db 839 QCNPRPKSLDIGAKGGSYDPRH 861

RESULT 14
Q8BX28 PRELIMINARY; PRT; 604 AA.
AC Q8BX28;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Weakly similar to N-acetylglucosamine-6-sulfatase (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).";
DR EMBL; AK049170; BAC33584.1; --
FT NON_TER 1
SQ SEQUENCE 604 AA; 69596 MW; A317924B658C59B7 CRC64;

Query Match 64.4%; Score 3059.5; DB 11; Length 604;
Best Local Similarity 92.7%; Pred. No. 1.6e-235;
Matches 561; Conservative 18; Mismatches 19; Indels 7; Gaps 3;

QY 272 HMEFTNMLQRLQTLMSVDDSMETIYNMLVETGELDNTYIVYADHGHIQFGLVKGK 331
Db 1 HMEFTNMLQRLQTLMSVDDSMETIYNMLVETGELDNTYIVYADHGHIQFGLVKGK 60
QY 332 SMPYEDIRVPFVVRGPNVEAGCLNPHVILNIDLAPTILDIAGLDIPADMDGKSLKLLD 391

```



```
Db 61 SMPYEFDIRVPFVYRGVNVAGSLNPHIVLNI DLAPTILDIAGLDIPADMDGKSLKLLID 120
QY 392 TERPVNRFHLKKKORVWRDQFLVERGKLHKRDNKVDQAQENFLPKYORVLDLCORAEY 451
Db 121 SERPVNRFHLKKLRVWRDQFLVERGKLHKRBDKVAQENFLPKYORVLDLCORAEY 180
QY 452 QTACEQLGQKQWQVEDATGKLKHLKCKGPMRL---GGSRLSNLVPKYTGCGSEACTCDS 508
Db 181 QTACEQLGQKQWQVEDASGTLKHLKCKGPMRFGGGGSRLSNLVPKYDQGSSEACSCDS 240
QY 509 ---CDYKLSLAGRKKLKKKKYKASYRSBSIRSAIEVDGRVYHVLGDAAQPRNLTKR 565
Db 241 GGGDYKLGLAGRR-KLUFKKKKTSTYARNRSIRSAIEVDGEIYHVGLDTPVPPRNLSKP 299
QY 566 HWPAPEDQDDQKDGDFSGTGGPLPDYSAANPIKVTHRCYILENDTVQCCLDYKSLQAWK 625
Db 300 HWPAPEDQDDQKDGSGTGGPLPDYSAANPIKVTHRCYILENDTVQCCLDYKSLQAWK 359
QY 626 DHKLHIDHEITLQNKIKNREVRGHLKKRPECDCHKISYHTQHKGRKRGSSLPFF 685
Db 360 DHKLHIDHEITLQNKIKNREVRGHLKKRPECDCHKRISYHSQHKGRKRGSSLPFF 419
QY 686 RKGLEQDKVWLLREQKKLRKLKRLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLGP 745
Db 420 RKGLEQDKVWLLREQKKLRKLKRLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLGP 479
QY 746 FCACTSANNTYMCWRTINETHNPLFCFPATGFLVFDLNTDYPQLMNAVNTLDRDLVQ 805
Db 480 FCACTSANNTYMCWRTINETHNPLFCFPATGFLVFDLNTDYPQLMNAVNTLDRDLVQ 539
QY 806 LHVLQMLRSCGKYQCNPRTRNDLGLKDGSGYEYQFORRKPMPKRPSSKSLGQLW 865
Db 540 LHVLQMLRSCGKYQCNPRTRNDLGLRDSGSGYEYQFORRKPMPKRPSSKSLGQLW 599
QY 866 EGWEG 870
Db 600 EGWEG 604
```

RESULT 15

```
Q9UPSS PRELIMINARY; PRT; 818 AA.
ID AC Q9UPSS
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein KIAA1077 (Fragment).
GN KIAA1077.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Brain;
RX MEDLINE=93397452; PubMed=10470851;
RA Kikuno R., Nagase T., Ishikawa K., Hirose M., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:197-205(1999).
DR EMBL; AB029000; BAA83029.1; -.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
DR PROSITE; PS00523; SULFATASE_1; 1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 818 AA; 95067 MW; 43470781E44B871 CRC64;
```

Query Match 64.4%; Score 3059; DB 4; Length 818;
Best Local Similarity 65.6%; Pred. No. 2.7e-235;

```
Matches 552; Conservative 112; Mismatches 129; Indels 48; Gaps 11;
QY 55 DVELGSMVWNNKTRIMEOGGTHFINAFVTTMCCPSRSSILTKGVVHNHTYNNENCS 114
Db 1 DVELGSLQVNNKTRKIMBHGGAFFINAFVTTMCCPSRSSMLTKGVVHNHTYNNENCS 60
QY 115 SPSSQAOHESRTFAVYLNSTGYRTAFPGKYLYNEYSYVPPGKMKWGLLKNSREYNTL 174
Db 61 SPSSQAOHESRTFAVYLNSTGYRTAFPGKYLYNEYSYVPPGKMKWGLLKNSREYNTL 120
QY 175 CRNGVKKHGSYKDYLTDLITNDSVSFFRSKMYPHRPVLYKVISHAAPGPDSDAPQ 234
Db 121 CRNGVKKHGSYKDYLTDLITNDSVSFFRSKMYPHRPVLYKVISHAAPGPDSDAPQ 180
QY 235 YSRLFNASQHIITPSYNYAPNDKIMRYTCTPMKPIHMEFTNMLQRLQTLMSVDDSM 294
Db 181 FSKLYPNASQHIITPSYNYAPNDKIMRYTCTPMKPIHMEFTNMLQRLQTLMSVDDSM 240
QY 295 ETIYNMVLTEGELDNTYIVYTADHGHIHQPGLVKGSMPYDFDIRVPPYVPGNVEAGC 354
Db 241 ERLYNMVLTEGELDNTYIVYTADHGHIHQPGLVKGSMPYDFDIRVPPYVPGNVEAGC 300
QY 355 LNPHIVLNIIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKQGVWVWDSFLV 414
Db 301 IVQIVLNIIDLAPTILDIAGLDIPADMDGKSLKLLDTERPVNRPHLKKQGVWVWDSFLV 360
QY 415 ERGKLLHKRDNKVDQAQENFLPKYORVLDLCORAEYOTACEOLGQKQWQVEDATGKLK 474
Db 361 ERGKLLHKRDNKVDQAQENFLPKYORVLDLCORAEYOTACEOLGQKQWQVEDATGKLK 420
QY 475 HKCKGPMRLGGS-ALSNLVPKYGGSEACTCDSGDYKLSLAGRR-KKLFKK----KY 527
Db 421 HKCKGPMRLGGS-ALSNLVPKYGGSEACTCDSGDYKLSLAGRR-KKLFKK----KY 480
QY 528 KASYVRSRSIRSAIEVDGRVYHVLGD-----AAQPRNLTKRHWPG--APEDQDDKDG 580
Db 481 KPRFVHTQTRSLSVFEGEIVDINLEEEELQVLQPRNIAKRHDEGHKGRDLQASSGG 540
QY 581 -----DFSGTGGLPDYSAANPIKVTHRCYILENDTVQCCLDYKSLQAWKHLHIDH 633
Db 541 NGRMLADSSNAVGGP-----TTVRVTHKCFILPNDISIHCEELYQSARAWKHKAYIDK 595
QY 634 EIRTQWIKNIREVRGHLKKRPECDCHKISYHTQHKGRKLRH--GSSLHPPFRKQLE 691
Db 596 EIEALQDKIKNIREVRGHLKKRPECDCHKISYHTQHKGRKLRH--GSSLHPPFRKQLE 655
QY 692 KD-KVWLLRE-QIKKKKLRKLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLGPFAC 749
Db 656 VDSKQLQFKENNRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRRKRR 715
QY 750 TSANNNTYMCWRTINETHNPLFCFPATGFLVFDLNTDYPQLMNAVNTLDRDLVQHWQ 809
Db 716 TSANNNTYMCWRTINETHNPLFCFPATGFLVFDLNTDYPQLMNAVNTLDRDLVQHWQ 775
QY 810 LMELRSCGKYQCNPRTRNDLGLKDGSGYEYQFORRKPMPKRPSSKSLGQLWEGE 869
Db 776 LMELRSCGKYQCNPRTRNDLGLKDGSGYEYQFORRKPMPKRPSSKSLGQLWEGE 817
QY 870 G 870
Db 818 G 818
```

Search completed: October 15, 2003, 13:09:20

Job time : 62 secs